The Thesis Committee for Sutthaporn Tripoppoom
Certifies that this is the approved version of the following thesis:

Assisted History Matching Workflow for Unconventional Reservoirs

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Kamy Sepehrnoori

Wei Yu

Assisted History Matching Workflow for Unconventional Reservoirs

by

Sutthaporn Tripoppoom

Thesis

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Engineering

The University of Texas at Austin

May 2019

Dedication

To my beloved parents, Sutisith and Supatsorn, and my siblings, Surisa and Suthaphong Tripoppoom for their unconditional love and encouragement.

Acknowledgements

Firstly, I would like to express my deepest gratitude to my advisor, Dr. Kamy Sepehrnoori, for his continuous supervision and encouragement during the past two years. His guidance in weekly meetings and welcoming discussions has steered my research into the right direction. More importantly, I am so honored to work under his supervision.

My sincerest appreciation goes to Dr. Wei Yu for his thorough guidance and dedication since the beginning. I am so grateful for his support, patience and trust in me to develop the knowledge on this research topic. His expertise and experience in unconventional reservoir simulation and fractures modeling have broadened my technical perspective. More importantly, I am so thankful for his time to review my thesis and to give me an insightful academic writing skill.

My appreciation extends to previous graduate students, Marut Wantawin and Silpakorn Dachanuwattana, who had built the strong foundation and did an excellent job in assisted history matching for unconventional reservoirs. I am so thankful for our research group students for knowledge sharing and technical discussion. It is my privilege to work with these brilliant people. Also, I would like to thank my UT PGE friends and Thai association of Austin for a wonderful life experience in the US.

Moreover, I would like to acknowledge PTT Exploration and Production Plc. (PTTEP) for granting me a scholarship to pursue a master's degree at the University of Texas at Austin. The technical knowledge and experience I have gained during the past two years will certainly be beneficial to my future career.

Finally, I am so grateful for my parents, family and friends in Thailand for their endless love and support. Everything I have accomplished would not be possible without them.

Abstract

Assisted History Matching Workflow for Unconventional Reservoirs

Sutthaporn Tripoppoom, M.S.E.

The University of Texas at Austin, 2019

Supervisor: Kamy Sepehrnoori

The information of fractures geometry and reservoir properties can be retrieved from the production data, which is always available at no additional cost. However, in unconventional reservoirs, it is insufficient to obtain only one realization because the non-uniqueness of history matching and subsurface uncertainties cannot be captured. Therefore, the objective of this study is to obtain multiple realizations in shale reservoirs by adopting Assisted History Matching (AHM).

We used multiple proxy-based Markov Chain Monte Carlo (MCMC) algorithm and Embedded Discrete Fracture Model (EDFM) to perform AHM. The reason is that MCMC has benefits of quantifying uncertainty without bias or being trapped in any local minima. Also, using MCMC with proxy model unlocks the limitation of an infeasible number of simulations required by a traditional MCMC algorithm. For fractures modeling, EDFM can mimic fractures flow behavior with a higher computational efficiency than a traditional local grid refinement (LGR) method and more accuracy than the continuum approach.

We applied the AHM workflow to actual shale gas wells. We found that the algorithm can find multiple history matching solutions and quantify the fractures and reservoir properties posterior distributions. Then, we predicted the production

probabilistically. Moreover, we investigated the performance of neural network (NN) and k-nearest neighbors (KNN) as a proxy model in the proxy-based MCMC algorithm. We found that NN performed better in term of accuracy than KNN but NN required twice running time of KNN. Lastly, we studied the effect of enhanced permeability area (EPA) and natural fractures existence on the history matching solutions and production forecast. We concluded that we would over-predict fracture geometries and properties and estimated ultimate recovery (EUR) if we assumed no EPA or no natural fractures even though they actually existed. The degree of over-prediction depends on fractures and reservoir properties, EPA and natural fractures properties, which can only be quantified after performing AHM.

The benefits from this study are that we can characterize fractures geometry, reservoir properties, and natural fractures in a probabilistic manner. These multiple realizations can be further used for a probabilistic production forecast, future fracturing design improvement, and infill well placement decision.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 MOTIVATION

In shale reservoirs, the permeability of organic-rich mudrocks is very low in the magnitude of nano Darcy compared with the magnitude of 0.1 md-1000 md in conventional reservoirs. This means the oil and gas production from shale reservoirs by a traditional technique as used in conventional reservoirs is not achievable or economically viable. Fortunately, with the advancements of horizontal-well drilling and multi-stage hydraulic fracturing in the recent decades, oil and gas production from shale reservoirs has been unlocked and emerged to fulfill energy demand. As a result, this shale reservoirs production inevitably became one of the major portions of oil and gas production worldwide.

As a number of hydraulically-fractured, shale-reservoir wells have been increasing in the past decades, production data, which is always available at no additional cost, can be beneficial to help characterize fracture geometry and reservoir properties. These fractures properties such as effective fracture half-length, fracture height and fracture conductivity are not always available by other methods. Although microseismic information sometimes is available, it does not provide the effective fracture geometry or fracture conductivity as we obtain from the production data. Besides, the matched model from history matching can be used further for production forecast. However, history matching is an inverse problem that the solutions are non-unique. Different sets of reservoir and fractures properties can mimic the same historical production data. In addition, traditionally, engineer performs a manual history matching which is a tedious and time-consuming task. Above all, after all these tedious tasks, only one realization of the history match solution can be obtained. This causes concerns that only one calibrated model is not capable to capture reservoir and fractures uncertainties and history matching non-

uniqueness. Furthermore, it is too risky for further production forecast, reserves estimation for asset valuation, further fracturing design improvement and well placement decision. Therefore, this is why Assisted History Matching (AHM) has been adopted to tackle this problem, which leads to the objective of this study.

1.2 Research Objectives

Because a traditional manual history matching is infeasible and has only one retrieved realization, we cannot capture subsurface uncertainties and the non-uniqueness of history matching. The objectives of this study are to obtain multiple realizations from history matching in shale reservoirs by performing Assisted History Matching (AHM) and then to perform production forecast probabilistically. The AHM tool that will be developed should be a workflow that is practical, efficient and accurate.

1.3 Thesis Outline

The thesis consists of seven chapters and is structured as follows. Chapter 2 reviews the current literature regarding assisted history matching in shale reservoirs. Different history matching algorithms along with their advantages and limitations are discussed. In addition, several fractures modeling methods are investigated and compared. Also, several challenges in shale reservoirs are presented and lastly, assisted history matching in unconventional reservoirs is discussed.

Chapter 3 emphasizes the proposed methodology of AHM workflow in shale reservoirs. The proposed AHM workflow is developed based on the philosophy of practicability, efficiency and accuracy. Firstly, the framework of AHM workflow is presented. Then, each component of the AHM framework is discussed in detail including Embedded Discrete Fracture Model (EDFM), reservoir simulator, proxy model (k-nearest

neighbors and neural network used in this study), multiple proxy-based Markov Chain Monte Carlo (MCMC) algorithm. Lastly, the steps in AHM workflow are explained.

Chapter 4 illustrates the application of the proposed AHM workflow to an actual shale gas well. The comparison between EDFM and local grid refinement (LGR) is firstly presented. Then, step-by-step in AHM workflow is emphasized including parameter identification and screening, history matching and production forecast. The multiple-proxy based MCMC algorithm using k-nearest neighbors is adopted in this chapter.

Chapter 5 continues to improve the history-matching algorithm by evaluating the comparison between neural network and k-nearest neighbors as proxy models in the proxy-based MCMC algorithm. The comparison of number of solutions found and elapsed time of both proxy models are demonstrated by an application of AHM to another shale gas well. Then, the effect of having enhanced permeability area (EPA) in fractures modeling to the history matching solutions and production forecast is investigated in the same shale gas well using neural network as proxy model. Two scenarios of the cases with and without EPA along with hydraulic fractures are performed.

Chapter 6 adopts the proposed AHM workflow to study the impact of natural fractures existence in shale gas reservoirs. Neural network as proxy model in proxy-based MCMC algorithm is utilized to perform AHM in shale gas well. Two scenarios of the case with and without natural fractures are presented. Natural fractures parameters are included as uncertain parameters in AHM for the case with natural fractures instead of using deterministic values.

Lastly, Chapter 7 summarizes the key findings of the study including their interpretation and implication, and the benefits of the study. Then, recommendations for future work are provided to improve the study in several aspects including the proxy-based

MCMC algorithm and AHM workflow codes, fractures modeling and the possible applications of the AHM workflow.

Chapter 2: Literature Review

The purpose of this chapter is to review and summarize the current literature of assisted history matching in unconventional reservoirs (shale reservoirs). We discuss the history matching algorithms along with their benefits and limitations, different fractures modeling techniques, the challenges in unconventional reservoirs and the assisted history matching in unconventional reservoirs.

2.1 History Matching Algorithm

There are many assisted history matching algorithms (AHM) but can be grouped into three categories including optimization-based methods, Ensemble-based methods and Bayesian Markov Chain Monte Carlo (MCMC) (Oliver and Chen 2011).

Optimization-based methods are popular for AHM to obtain multiple realizations due to their efficiency and relatively low computational efforts. Some algorithms rely on gradient or derivative of objective function calculation such as gradient method (Ding 2011) and Gauss Newton method (Gao et al. 2017; Chen et al. 2017). Others use the derivative-free method including particle swarm (Vazquez et al. 2015) and evolutionary algorithm (Yin et al. 2011; Zhang et al. 2013; Mirzabozorg et al. 2013; Xie et al. 2015). However, the optimization-based methods are not suitable for effectively quantifying uncertainty as history matching problem. Moreover, the history matching solutions can be biased by using different optimization algorithms (Goodwin 2015). Erbas and Christie (2007) also demonstrated that different optimized algorithms (genetic algorithm and neighborhood algorithm) affect the history matching solution ensemble and production forecast. Ensemble-based methods such as Ensemble Kalman Filter (EnKF) (Emerick and Reynolds 2011) and Ensemble Smoother (EnS) (Chai et al. 2018, Chang and Zhang 2018) are another AHM method using covariance matrix to update ensemble of parameters.

Nevertheless, in most cases, history matching is a highly nonlinear problem, which is challenging for Ensemble-based method assumption of linear relationship between input and output parameters (Emerick and Reynolds 2012; Goodwin 2015).

Therefore, the method in Bayesian inference as MCMC, including several types such as Gibbs, Metropolis-Hasting (MH) and Hamiltonian, is used for history matching to capture history-matching uncertainties. The reason is that MCMC has benefits of quantifying uncertainty without bias or being trapped in any local minima. However, due to the fact that MCMC is a sampling algorithm normally requiring more than $105\text{-}10^6$ steps to reach a stable final distribution, there is a limitation of computation effort for the traditional MCMC requiring infeasible number of simulations. Many studies have been done to overcome this issue. Maucec et al. (2007) used a semi-analytical streamline model together with MCMC for history matching. Elsakout et al. (2015) proposed multilevel MCMC to use coarser grid before finer grid with MCMC. However, one prominent solution is to use surrogate or proxy model to minimize the expensive reservoir simulation runs of those MCMC steps (Slotte and Smorgrav 2008; Goodwin 2015; Yeh et al. 2016). Various proxies are used together with MCMC as AHM algorithm. Wantawin et al. (2017a, 2017b) used polynomial proxy with MCMC to perform history matching in shale gas wells. Yu et al. (2018c) and Tripoppoom et al. (2019) used k-nearest neighbor (KNN) proxy with MCMC for history matching and probabilistic production forecast of shale gas wells.

Dachanuwattana et al. (2019) also made comparison of using different proxies including polynomial, KNN and Kriging and found that Kriging is relatively more predictive proxy than the rest while it needs more computational efforts than others as well. They also found that KNN could be a balance choice between model predictability and computational effort. Indeed, machine learning such as neural network (NN) can be

beneficial as the proxy model. Some studies used NN with other history matching applications such as NN with optimization-based method (Cullick et al. 2006), NN with pattern recognition technology for permeability field history matching (Shahkarami et al. 2015) and NN in proxy-based Acceptance-Rejection algorithm (PAR) (Yang et al. 2015). Due to the flexibility of NN formulation and training process to mitigate overfitting issues, NN could fit any highly non-linear function better than KNN proxy. Nevertheless, there are very few studies utilizing a machine learning technique such as NN together with MCMC algorithm for AHM workflow.

To sum up, to achieve the practical application, and efficient and accurate history matching results, one of the most suitable algorithms to perform assisted history matching should be in the category of Bayesian MCMC together with proxy models to solve the issue of an infeasible number of simulation runs of a traditional MCMC.

2.2 Fracture Modeling

Another aspect to be considered for assisted history matching in shale reservoirs is fracture modeling for both hydraulic fractures and natural fractures.

There are several modeling techniques to mimic fractures flow behavior. One traditional method is a continuum approach such as dual porosity dual permeability model (DPDK). However, the continuum approach has the limitation to model non-uniform and irregular spaced fractures due to its assumption of uniform distributed fractures. Rubin (2010) also found that the continuum approach cannot accurately capture the fracture-matrix connection in very low permeability reservoirs. This leads to the advent of Discrete Fracture Model (DFM). DFM using Local Grid Refinement (LGR) (Cipolla et al. 2009) has been used to accurately model fractures but one major issue is that it requires tremendous computational effort especially when the number of fractures is high, and

fracture becomes more complex. Methods such as logarithmically spaced, locally refined and dual permeability grid (LS-LR-DK) by Rubin (2010) also tried to improve model accuracy but it is still around the concept of LGR. Therefore, a general 3D Embedded Discrete Fracture Model (EDFM) has been developed to combine benefits from the continuum approach and LGR (Moinfar et al. 2014; Cavalcante Filho et al. 2015a; Shakiba et al. 2015; Xu et al. 2017a, 2017b; Yu et al. 2018b, 2018d). The EDFM can model fractures flow behavior with an acceptable accuracy as LGR but with much less computational efforts (Moinfar et al. 2014; Xu et al. 2017a, 2017b, 2018; Yu et al. 2018d). Many applications of EDFM have been done due to its computational efficiency such as interwell fracture interference (Fiallos et al. 2019a, 2019b) and CO2 Huff-n-Puff (Yu et al. 2019b).

To model fractures, in conventional reservoirs, we can model the hydraulic fractures with the bi-wing planar type, which is acceptable. While for shale reservoirs, in some cases, the bi-wing planar type assumption sometimes cannot represent the actual hydraulic fracture propagation because the hydraulic fractures can interact with pre-existing natural fractures and will induce the complex fracture network rather than bi-wing planar fractures. Figure 2.1 (Warpinsky et al. 2008) shows different fracture complexity from simple to extremely complex. More and more results from microseismic also support the complex induced fracture network in shale reservoirs as illustrated in Figure 2.2 (Fisher et al. 2004).

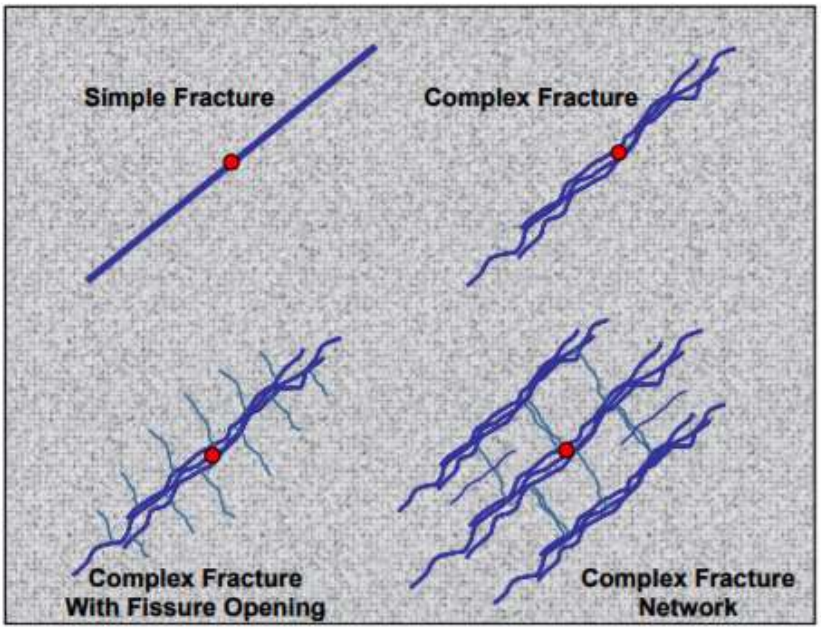


Figure 2.1: Different types of fracture complexity from simple fracture to complex fracture network. (Ref. Warpinsky et al. 2008).

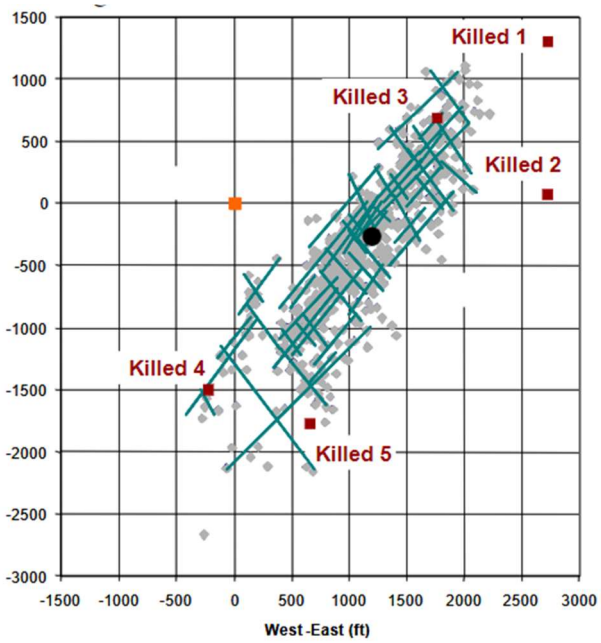


Figure 2.2: Microseismic events map in vertical well showing complex fracture network in Barnett. (Ref. Fisher et al. 2004).

To obtain the complex fracture networks, a geomechanical model is utilized to predict the fracture propagation (Cipolla et al. 2012; Taleghani et al. 2013; Wu and Olson 2015; Shrivastava et al. 2018) considering reservoir properties, fracturing operation and the interaction with pre-existing natural fractures. However, the complex fracture geometries generated from this fracture propagation model are a challenging task to obtain accurate and reliable results. This is because there are high uncertainties and assumptions in the model and input information is sometimes limited. Therefore, many studies attempted to compromise the concept of complex fractures by a simpler concept of simulated rock volume (SRV) rather than the fully complex fractures from the geomechanical model. The SRV is defined as the stimulated reservoir volume occurred by hydraulic fractures and their interaction with pre-existing fractures (Mayerhofer et al. 2010).

Du et al. (2010) adopted the dual porosity model to represent fracture networks in shale gas well. Chang and Zhang (2018) used dual permeability and dual porosity to model SRV together with the explicit hydraulic fractures for history matching in shale gas reservoirs. Besides using the continuum approach alone, Rubin (2010) modeled the SRV area by the marriage of LGR concept and dual permeability model, or known as logarithmically spaced, locally refined, dual permeability grid (LS-LR-DK). Yin et al. (2011) used another method of enhanced permeability area together with explicit hydraulic fractures to represent the SRV region. Shakiba and Sepehrnoori (2015) used EDFM to model fracture networks.

To estimate SRV, fracture extent and orientation, several studies utilized the outputs from microseismic interpretation. Xie et al. (2015) utilized microseismic information and wiremesh discrete fracture models to represent SRV from microseismic. They also integrated analytical solution of drainage volume calculation by fast marching method with

genetic algorithm to perform history matching of production data. Shakiba and Sepehrnoori (2015) used microseismic to constrain the SRV modeling by using EDFM, which provides more running computational efficiency than a traditional LGR.

Other studies also utilized microseismic interpretation to calibrate geomechanical model. Shaffner and Cheng et al. (2011) used microseismic temporal and spatial events to confirm the completion diversion during each fracturing stages. They also used microseismic events mapping for fractures geometry over time to calibrate the fracture mechanical model such as fracture half-length, fracture height, and azimuth. Xu et al. (2009) integrated microseismic with semi-analytical pseudo 3D-geomechanical model. They considered the mass conservation of injected fluid and mechanic behavior interaction between fluid and fractures. The spatial and temporal microseismic events were utilized to calibrate the fracture propagation, fluid pressure and fracture width in the fracture model.

Nevertheless, microseismic is not always available in every well. Chang and Zhang (2018) show that history match of SRV without microseismic interpretation is a challenging task. They investigated the non-uniqueness of SRV shape and volume by performing production history matching. This will highlight the benefits of having microseismic as the initial constrain of fracture geometry and for fracture model calibration.

In addition, one limitation of microseismic is that it cannot provide the effective fractures that contribute to the production. In other words, the fracture conductivity and propped/unpropped fracture area cannot be extracted from microseismic (Maxwell et al. 2013). This causes the fracture geometry after calibrating with the actual production history or history matching always less than the fracture geometry from microseismic (Sharma et al. 2004). Cipolla and Wallace (2014) also emphasized that SRV provided from

microseismic may not be a good indication of the effectiveness of fractures. They also supported the using of fractures modeling to estimate the propped and un-propped fracture area by performing production history matching. Mayerhofer et al. (2010) also used reservoir and fracture model to evaluate the effective fractures by varying fracture spacing for many cases with identical SRV from microseismic. Shakiba et al. (2015) also demonstrated microseismic cannot capture the small fissures due to microseismic resolution because microseismic can only capture 0.1% of the fracturing events, which most of them are aseismic process.

In sum, the main limitation of using microseismic for fracture modeling is that microseismic cannot give any further information about the effective fractures volume. If engineers only calibrated the fracture mechanical model or used fracture modeling techniques such as DPDK or discrete fracture geometry from only microseismic interpretation, the simulation results will not be accurate. Also, not every well has microseismic information. Therefore, to determine the effective SRV, or effective fracture geometry, the need to calibrate with production historical data, which is always available at no additional cost, is crucial for later production forecast or fracturing design and future well spacing placement.

2.3 Other Challenges in Unconventional Reservoirs

To perform AHM in unconventional reservoirs such as shale reservoirs, there are many challenges in addition to those in conventional reservoirs. Those challenges include complicated fluid transport in nanopores and reservoir heterogeneity and natural fractures existence.

Firstly, the complex fluid transports in nanopores are illustrated by Javadpour et al. (2007, 2009) and Shabro et al. (2011, 2012). Those complicated mechanisms are such as

slip flow, Knudsen diffusion, gas desorption and gas diffusion in kerogen, which the traditional Darcy's law cannot be directly applied. Secondly, in shale reservoirs, very high heterogeneity in reservoir properties are normally observed as demonstrated by Aderibigbe et al. (2016) and Amin et al. (2016). They performed multi-mineral analysis using well logging information and the results indicated the high heterogeneity in a vertical direction. Different rock qualities could be observed and is an important criterion to be selected as production target. Also, Yu et al. (2018b) demonstrated the effect of gas desorption, gas slippage and diffusion, non-Darcy flow to production forecast in shale gas reservoirs. In the case study, the total effect could increase 30- year gas recovery by 18%.

Another challenge in unconventional reservoirs is the existence of natural fractures in shale reservoirs. Natural fractures not only affect fractures propagation during hydraulic fracturing (Taleghani et al. 2014; Wu and Olson 2016; Agrawal et al. 2019) but also have an effect to the production performance in the long term. Yu et al. (2018b) investigated the effect of natural fractures existence modeled by EDFM and found that the gas recovery could be increase by 23.2% in the case study when natural fractures exists. Moreover, Yu et al. (2019a) showed the impact of natural fractures on the long-term production forecast of hydraulic-fractured well and well spacing decision. Many studies attempted to characterize natural fractures using information from core data, image log, seismic and outcrop. Shrivastava et al. (2018) used core data to characterize natural fractures and calibrated fracture propagation model. However, the information about natural fractures is challenging to obtain, and often is not available. On the other hand, production data, which is always available, can be used to help reducing natural fractures uncertainty by performing assisted history matching. By using EDFM for fractures modeling, a higher computational efficiency of modeling natural fractures discretely can be achieved during the assisted history matching.

2.4 Assisted History Matching in Unconventional Reservoirs

Due to these challenges in shale reservoirs, there are several studies attempting to understand more on shale reservoirs. For instance, analytical models such as Rate Transient Analysis (RTA) have been used to characterize shale reservoirs and forecast production with various assumptions (Guo et al. 1994; Raghavan et al. 1997; Brown et al. 2009; Stalgorova and Mattar 2013). Luo et al. (2014) combined RTA with Monte Carlo analysis to evaluate uncertainty in matrix and fracture properties leading to uncertainty in production forecast. Even though RTA is a fast method to characterize reservoir and forecast the production, the limitation of RTA is that there are many simplified assumptions about fluid type and hydraulic fractures pattern and heterogeneity, which cause the constraint in application to the real complex unconventional fields. Moinfar et al. (2016) demonstrated the limitation of RTA assumptions applied to real field and emphasized the benefits of using numerical model over analytical model.

With the limitation of RTA, this leads to several studies on numerical models in unconventional reservoirs (Cipolla et al. 2009; Rubin 2010; Novlesky et al. 2011; Yu and Sepehrnoori 2014, 2018; Yu et al. 2016, 2018a). The study of numerical model is later extended to integrate with AHM in unconventional reservoirs. Nejadi et al. (2015) used discrete fracture network with upscaling technique and EnKF for the shale gas well in Horn River, Canada. Yang et al. (2015) adopted LGR to model fractures and the proxy-based Acceptance-Rejection method was applied for HM in shale oil reservoir. Wantawin et al. (2017a, 2017b) used LGR to model fractures and polynomial proxy with MCMC was used for an iterative AHM in shale gas wells.

However, LGR is computationally expensive especially in the complex fractures scenario. The using of EDFM with AHM can improve the efficiency of history matching

Nevertheless, a few implementations of EDFM with AHM in unconventional reservoirs have been performed. Cavalcante Filho et al. (2015b) used EDFM and Simplex optimization algorithm for AHM in fishbones wells. Eltahan et al. (2019) coupled EDFM with commercial AHM tool to perform history matching in shale oil well. Dachanuwattana et al. (2018a, 2018b) used proxy-based MCMC algorithm with EDFM in a shale oil and shale gas condensate well. Yu et al. (2018c) and Tripoppoom et al. (2019) used multiple-proxy based MCMC algorithm with EDFM for shale gas reservoir.

To conclude, by considering all advantages and disadvantages of each method in the aspects of history matching algorithm, fractures modeling and challenges in unconventional reservoirs, the use of proxy-based MCMC algorithm with EDFM and numerical reservoir simulator could be a practical, efficient and accurate AHM workflow for unconventional reservoirs. This could achieve the goal of the study of obtaining multiple history matching solutions and performing the production forecast probabilistically.

Chapter 3: Methodology

3.1 Introduction

This chapter will discuss the methodology we are using to perform Assisted History Matching (AHM) in shale reservoirs. The methodology we used revolves the concept of a practical, efficient and accurate workflow. We start explaining what components consist in a framework of the AHM workflow to achieve the objective of the study. Then, each element in the framework is discussed including Embedded Discrete Fracture Model (EDFM) preprocessor, reservoir simulator, proxy model and proxy-based Markov chain Monte Carlo (MCMC) algorithm. Finally, the steps in AHM workflow are summarized in flow charts.

3.2 Assisted History Matching (AHM) Framework

The AHM framework consists of four components including EDFM preprocessor, reservoir simulator, proxy model and proxy based-MCMC algorithm, as shown in Figure 3.1. Firstly, the fractures are modelled by EDFM preprocessor to pre-calculate the non-neighboring connections. Then, the outputs from EDFM preprocessor are coupled to commercial simulator DATA file. Next, the proxy model is constructed to be used for the history-matching (HM) algorithm. The proxy-based MCMC algorithm is used as history matching algorithm in the AHM workflow. This workflow continues automatically until the proxy model is converged or the maximum number of simulations is exceeded.

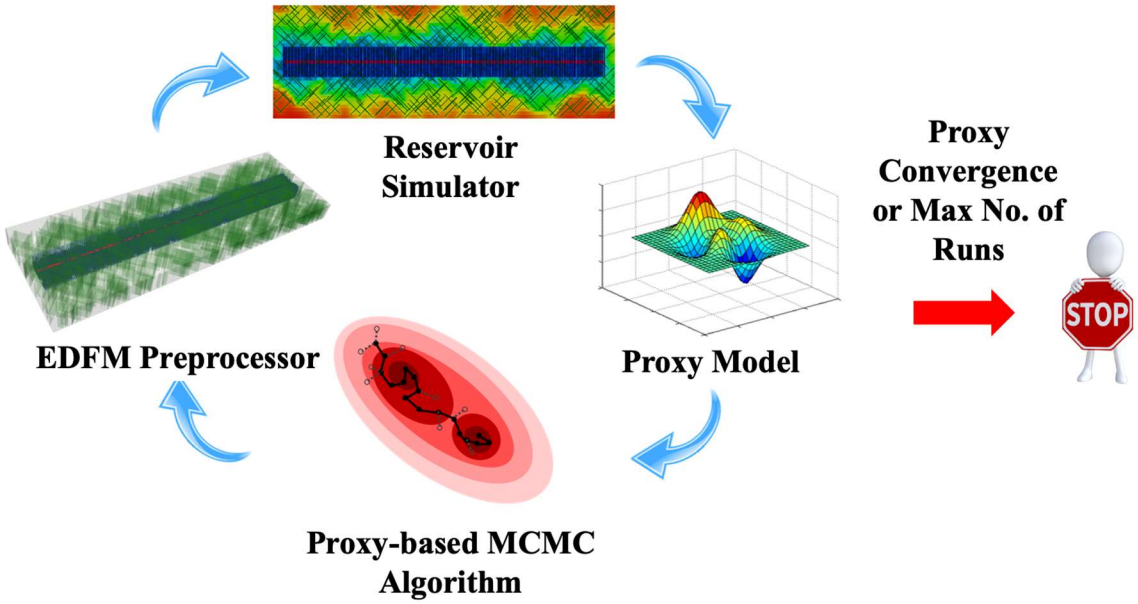


Figure 3.1: Integrated framework of Assisted History Matching (AHM) workflow.

3.3 Embedded Discrete Fracture Model (EDFM)

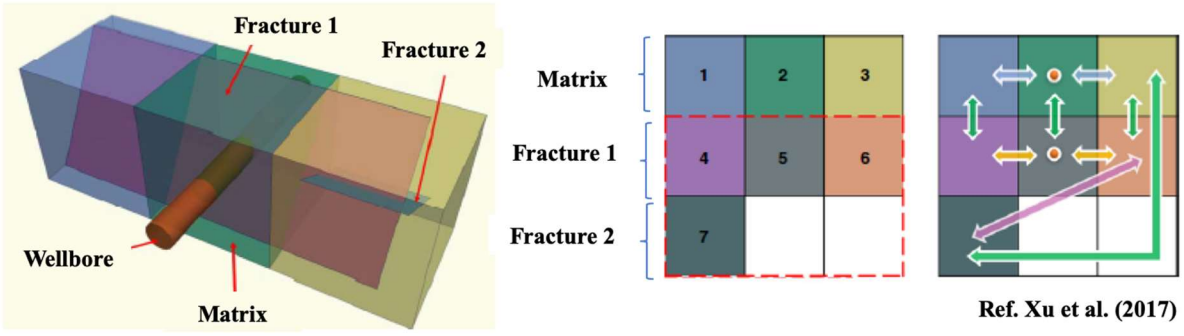


Figure 3.2: Concept of Embedded Discrete Fracture Model (EDFM)

The EDFM method combines the benefits from both the continuum approach and local grid refinement (LGR) method. As shown in Figure 3.2, an example with three matrix

grid blocks, two fractures and wellbore, the EDFM concept is that it creates two sets of grids including matrix and fractures. The fractures grid blocks are generated by the intersection between fractures and matrix grid blocks. Then, each grid block communicates through non-neighboring connection keywords.

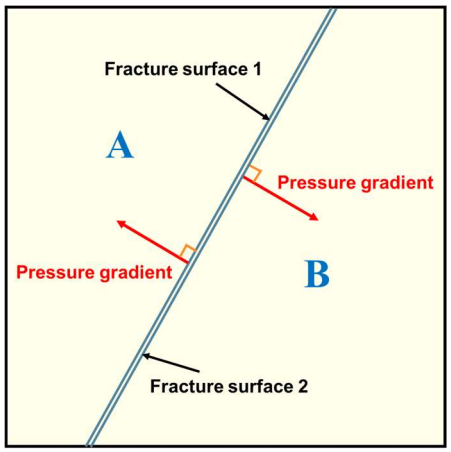
Because fractures are superimposed into matrix grid and linear pressure gradient is assumed, there is no need to discretize a lot of small grid blocks as the traditional LGR. With this lower number of grid blocks, EDFM could provide an acceptable accuracy as LGR but with lower computational time than the LGR. This improvement becomes more dominant when fracture networks are more complex (Xu et al. 2017a, 2017b, 2018).

As illustrated by Xu et al. (2017a), EDFM creates the additional cells to represent fractures. Each additional cell is assigned to each fracture segment discretized by matrix cell boundaries. The effective porosity for each additional fracture cell, ϕ_f, is calculated by

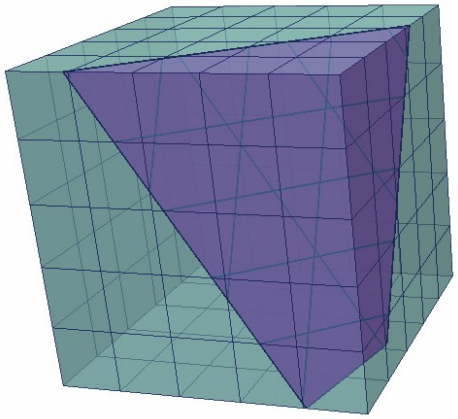
$$\phi_f = \frac{S_{seg} w_f}{V_b}, \tag{3.1}$$

where S_{seg} is the area of the fracture segment perpendicular to the fracture aperture , w_f is the fracture width and V_b is the bulk volume of the cell assigned for the fracture segment.

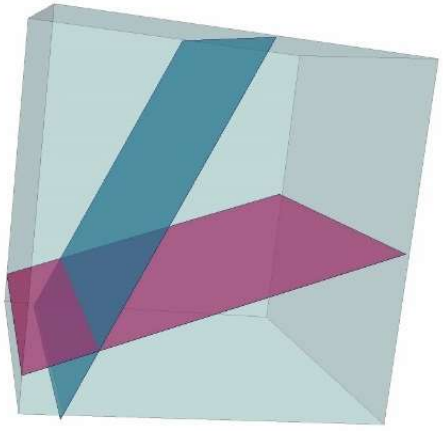
Then, non-neighboring connections (NNCs) are later defined to allow the transport phenomena between cells. There are four types of NNCs including matrix-fracture, fracture-fracture in an individual fracture, fracture-fracture of two different fractures and well-fracture as shown in Figure 3.3.



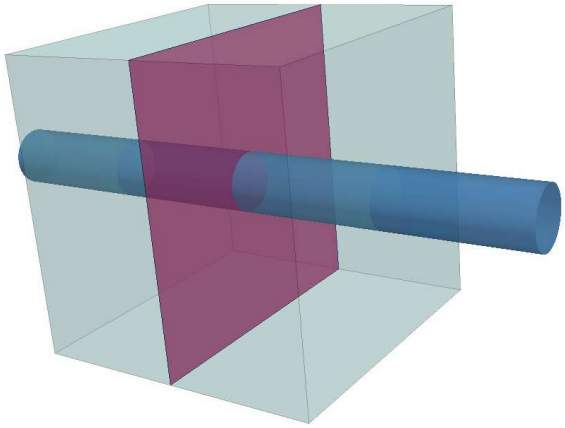
(a) Fluid flow between matrix grid blocks and fracture segments



(b) Fluid flow between fracture segments within an individual fracture



(c) Fluid flow between intersecting fracture segments within different fractures



(d) Fluid flow between fracture segments and wellbore

Figure 3.3: Four types of NNC in the EDFM method (Xu, 2015; Xu et al., 2017a).

Xu et al. (2017a) explained the calculation of NNCs in detail but in short, each pair of cells is communicated by transmissibility factors. The volume flow rate of phase l between two cells in an NNC pair is calculated by

$$q = \lambda_l T_{NNC} \Delta P, \tag{3.2}$$

where λ_l is the relative mobility of phase *l*, T_{NNC} is the NNC transmissibility factor, ΔP is the potential difference between two points. The transmissibility calculation is different depending on what type of NNCs. For the matrix-fracture connection, the NNC transmissibility depends on the matrix permeability and fracture geometry and is calculated by

$$T_{NNC} = \frac{k_{NNC} A_{NNC}}{d_{NNC}}, \tag{3.3}$$

where k_{NNC} is the matrix permeability in the normal direction to the fracture plane, A_{NNC} is the contact area of the fracture plane inside the matrix block and d_{NNC} is the average normal distance from matrix block to fracture plane. For the fracture-fracture connection, the NNC transmissibility is calculated by using two-point flux approximation scheme:

$$T_{NNC} = \frac{T_1 T_2}{T_1 + T_2}, \tag{3.4}$$

The detailed formulations to calculate each fracture segment transmissibility (T_1 and T_2) depend on either the two fracture segments are in the same individual fracture discretized by matrix cell boundary or they are from two different fractures. For the well-fracture connection, the effective well index is assigned to model each fracture segment that intersects the well trajectory. The well index formulation, WI_f, (Moinfar et al. 2013), is derived from the formula proposed by Peaceman (1983):

$$WI_f = \frac{2\pi k_f w_f}{\ln\left(\frac{r_e}{r_w}\right)}, \tag{3.5}$$

$$r_e = 0.14\sqrt{L_s^2 + H_s^{\,2}}, \tag{3.6}$$

where k_f is the fracture permeability, r_e is the effective wellbore radius, r_w is the wellbore radius, L_s is the length of the fracture segment, and H_s is the height of the fracture segment.

3.4 COMMERCIAL RESERVOIR SIMULATOR

For this study, after the outputs from EDFM preprocessor are obtained, these outputs will be coupled with a commercial reservoir simulator. CMG-IMEX for shale oil and CMG-GEM for shale gas reservoirs are used in this study. Our developed code in the proposed AHM workflow will generate the input to EDFM preprocessor and output to edit reservoir simulation data file automatically.

3.4.1 Multiphase Flow Governing Equations

For reservoir simulator, the governing equations is 3D multiphase flow black oil model. Starting with a mass conservation, mass in minus mass out equals to net change in mass content. The continuity equation is shown in Equation 3.7. Then, Darcy's law of Equation 3.8 shows the relationship between the potential difference and velocity of phase "f". With accounting for the well, from continuity equation and Darcy's law, they become flow equations as shown in Equations 3.9 and 3.10 (Aziz and Settari 1979). In addition, we have auxiliary equations of capillary pressure to relate each phase pressure, and saturation of all phases must be equal to 1, as shown in Equations 3.11 and 3.13.

Continuity Equation

$$-\nabla \cdot (\rho_f \vec{v}_f) = \frac{\partial}{\partial t}(\emptyset s_f \rho_f) \tag{3.7}$$

Darcy's Law

$$v_{fs} = -\left(\frac{k_s k_{rf}}{\mu_f}\right)\frac{\partial \Phi}{\partial t} \tag{3.8}$$

Flow Equations

$$\frac{\partial}{\partial x}\left(\frac{A_x k_x k_{rf} \partial \Phi_f}{\mu B_f \quad \partial x}\right)\Delta x + \frac{\partial}{\partial y}\left(\frac{A_y k_y k_{rf} \partial \Phi_f}{\mu B_f \quad \partial y}\right)\Delta y + \frac{\partial}{\partial z}\left(\frac{A_z k_z k_{rf} \partial \Phi_f}{\mu B_f \quad \partial z}\right)\Delta z + \; q_{fsc} = \frac{v_b}{5.615}\frac{\partial}{\partial t}\left(\frac{\emptyset s_f}{B_f}\right) \tag{3.9}$$

$$q_{fsc} = -PI\left(\Phi_{block} - \Phi_{sf,bottomhole}\right) \tag{3.10}$$

where ρ_f is density of phase f, $\vec{v}_f$ is velocity vector of phase f, $\emptyset$ is posity, s_f is saturation of phase f, k_s is permeability, k_{rf} is relative permeability of phase f, μ_f is viscosity of phase f, Φ is potential equal to P_f - $\rho_f g$, A is area, B_f is formation volume factor of phase f, v_b is bulk volume, q_{fsc} is flow rate of phase f at standard condition, and PI is productivity index.

Auxiliary Equations

$$P_{cow}(S_w) = P_o - P_w \tag{3.11}$$

$$P_{cgo}(S_g) = P_g - P_o \tag{3.12}$$

$$S_o + S_w + S_g = 1 \tag{3.13}$$

where P_{cow} is capillary pressure between oil and water, and P_{cgo} is capillary pressure between gas and oil.

The boundary and initial conditions for this study problem are as follows:

- Neumann boundary condition
    - No flow boundary: $\nabla P = 0$
- Initial condition
    - Initial pressure (P) = reservoir pressure
    - Other parameters as initial condition input are history matching parameters or uncertain parameters.

The simulation is constrained by either bottomhole pressure or main fluid flow rate (oil flow rate, gas flow rate).

3.4.2 CMG Reservoir Simulator

The summary of CMG reservoir simulator is as below,

- Body-Centered Finite Difference Method.
- Fully Implicit Method to solve time dependent problem.
- Preconditioner order: REDBLACK (similar to D4 ordering) for ILU(1)
- CMG's AIMSOL (Non-parallel iterative solver)
- Residual check, material balance check

3.5 PROXY MODEL

Proxy model is a fast approximation of the original or actual model, which is normally more complicated and time-consuming to obtain the value. Depending on the purposes, proxy model is interchangeably used with several other terms such as surrogate model and response surface model, but the main objective of each term is the same. In this

study, two types of proxy models are utilized including K-Nearest Neighbors (KNN) and Neural Network (NN).

3.5.1 K-Nearest Neighbors (KNN)

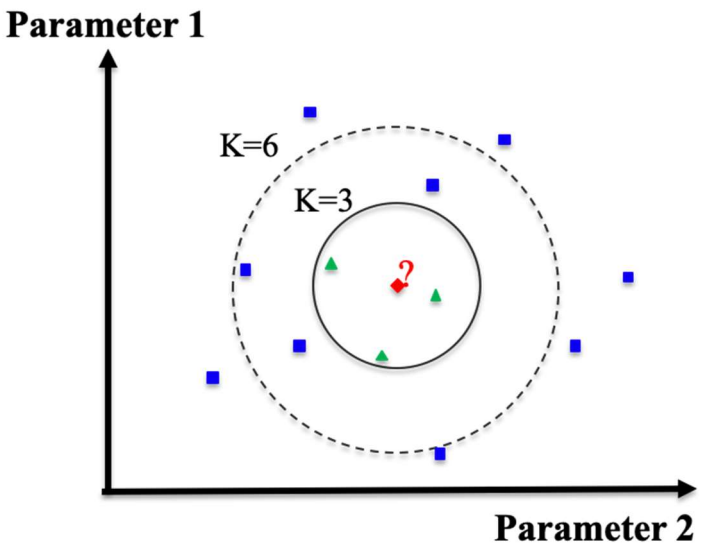


Figure 3.4: Concept of K-Nearest Neighbors (KNN)

The concept of KNN is to estimate the unmeasured point value by using information from measured neighboring points. As shown in Figure 3.4, the predicted value is a function of two parameters, and we would like to predict the value at red point by having available measured points around the red point. However, the problem is how many neighboring points should be used. To answer this question, K-fold cross validation is performed, evaluating from K=1 to 1.5 times number of measured points. For example, there are 50 measured points. These 50 points are generated into 5 folds. Each fold consists of different set of 40 training points and 10 validating points. The nearest K points for each response-parameter proxy is selected based on one giving the lowest root mean square

error; in other words, the highest R squared. As illustrated in Figure 3.4, if K of 3 gives the lowest lower, the red point values will be estimated from the nearest 3 points surrounding it as shown as green points. To construct KNN proxy model, each uncertain parameter is normalized to eliminate the effect of different value ranges. Then, the value of unmeasured point is estimated by averaging the values of k- nearest measured points and giving more weight to the nearer points as Equation 3.14. The inverse distance weighted averaging method is used.

$$\hat{z}(\vec{\theta}_0) = \sum_{i=1}^{k} \lambda_{i0} z(\vec{\theta}_i) \quad , \tag{3.14}$$

where $\vec{\theta}_0$ is a vector of independent variables (uncertain parameters), $\hat{z}(\vec{\theta}_0)$ is the estimated response (response parameters of unmeasured points), $\vec{\theta}_i$ is a vector of the nearest K measured points, $z(\vec{\theta}_i)$ is the response of measured points, and λ_{i0} is the weight function between $\vec{\theta}_0$ and $\vec{\theta}_i$ defined by Equation 3.15.

$$\lambda_{i0} = \frac{\sum_{i=1}^{k} \left| \vec{\theta}_i - \vec{\theta}_0 \right|}{\left| \vec{\theta}_i - \vec{\theta}_0 \right|} \tag{3.15}$$

There are several benefits to use KNN as a proxy model. Firstly, KNN is an exact proxy and this means the value at measured point is an exact unlike polynomial fitting which needs to fit the polynomial curve to measured points and overfitting issue can exist. Secondly, KNN requires relatively low computational efforts. Even though Kriging provides more accurate estimated response based on variogram, it needs a lot more

computational time than KNN. As a result, KNN could be an acceptable proxy having the balance between accuracy and computational time.

3.5.2 Neural Network (NN)

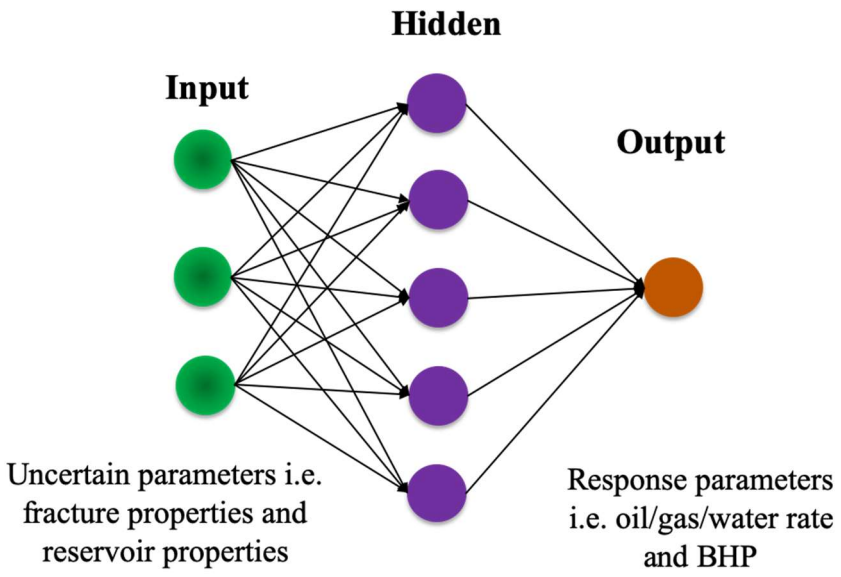


Figure 3.5: Structure of Neural Network consisting of input, hidden, and output layers.

Neural network is an artificial intelligence to mimic human neurons in the brain. The structure of neural network consists of input, hidden and output layers as shown in Figure 3.5. Between each layer, there are weight values assigned to each connection between each neuron. In each neuron, the combined input signals are calculated and compared to threshold level as the activation function for the output calculation. The structure of NN used in this study is fully connected as shown in Figure 3.5. The number of input parameters is equal to the number of uncertain parameters and the output for each NN has only one parameter for each response parameter. In this study, multiple proxies are utilized to handle non-linearity of the problem, so that several NNs were constructed for each

response parameter. To train the networks' weight values, the supervised training algorithm is adopted. In this study, Levenberg-Marquardt algorithm (Hagan and Menhaj, 1994) is used for back propagation feedback algorithm to train the neural networks. The Levenberg-Marquardt is similar to quasi-newton method that no Hessian matrix computation is needed while the algorithm maintains approximately the second order of convergence as Newton's method.

While NN is considered as a universal function that can fit any highly nonlinear behavior, one major concern of NN is about an overfitting of the network. In other words, neural network can fit the training data set accurately but it cannot generalize and predict the values when being applied with other data sets. To overcome this issue, the data set is divided into three separate groups including training, calibration and blind-test sets (Mohaghegh, 2017). The neural networks are firstly trained by each epoch of the data training set until no further improvement is observed, then the calibration data set, which has not been trained by the networks, is validated for the neural network prediction error. The objective of having calibration data set is to prevent the overtraining of neural networks. The training process will continue if both training and calibration data sets are still improving. However, the networks will stop training when no improvement of prediction error in calibration data set is observed. Therefore, neural network is considered as converged. With this manner, the overfitting issue can be mitigated. For the blind-test data set, this is used as a tracker and it has no effect to the neural network training process. For the portion of each data set, more portions of data sets should be partitioned to the training data set and less portion assigned to the calibration and blind-test data sets. In this study, 70:15:15 was used for the training, calibration and blind-test data sets, respectively.

There are several aspects to be considered for building the successful and predictive neural networks. Those aspects include how many hidden layers and neurons in each hidden layer should be used, what activation function should be adopted, and what portion of the training, calibration and blind-test data set should be assigned (Mohaghegh, 2017). However, these aspects are not the key factor to determine the success or failure of neural networks. Indeed, more important factor is the input parameters to train the networks. To illustrate, what parameters mainly affect the response parameters and should be used as the input, and what their ranges based on prior knowledge of this well and reservoir. Therefore, we should focus more on the quality of input information rather than other aspects of neural network construction.

3.6 PROXY-BASED MARKOV CHAIN MONTE CARLO (MCMC) ALGORITHM

3.6.1 Markov Chain Monte Carlo (MCMC)

The fundamental concept of all history matching algorithms is to find the parameter domain that can provide low objective function value or history matching solutions. Nevertheless, not all methods such as optimization-based method are for uncertainty quantification. MCMC, one of Bayesian inference methods, is the suitable method to quantify uncertainty of history matching. While a traditional Monte Carlo (MC) is a random sampling algorithm, MCMC could be viewed as the MC algorithm with memory. This means the next sampling point is proposed around the current point and decide whether to move to the next point based on the current condition. This is a reason why this method is called Markov Chain because the chain of sampling points is formed by the algorithm. In this study, Metropolis-Hasting (MH) MCMC is adopted. The MH MCMC, proposed by Metropolis et al. (1953) and Hastings (1970), is a random walk with probability. The concept is that if the next step gives lower objective function value than

the current state, there is higher probability close to one that the algorithm will step to the next point rather than staying at the current state. On the other hand, if the next step gives higher objective function value than the current state, there is lower probability closer to zero that the algorithm will step to the next point and may prefer staying in the current state. With this manner, after the sufficient step numbers (normally 10^5-10^6 steps), the MCMC algorithm will reach the stable distribution. Besides, with walking with probability concept, this algorithm will not be trapped in any local minima and can explore the entire domain.

The formulation and detailed steps of MH MCMC are summarized in Figure 3.6. Firstly, the Markov chain is initialized at the realization that has the lowest objective function. Then, the proposed uncertain parameters realization, θ^*, for the next step of Markov chain is calculated from a random walking distance, not exceeding the specified maximum distance from the previous points as Equation 3.16.

$$\theta^* = \theta + \delta, \tag{3.16}$$

where θ is the uncertain parameter realization at the current step and δ is the walking distance.

Then, the global error is estimated from the proxy. Normally, without using proxy together with MCMC algorithm, the simulation run is needed in order to get the global error value of the proposed case. However, with using proxy-based MCMC algorithm, the global error can be estimated from proxy model. Next, the acceptance probability ratio, α^*, is calculated, defined by Equation 3.17 with the assumption of Gaussian distribution.

$$\alpha^* = \min\left\{1, \frac{p\left(\theta^* \middle| x\right)}{p\left(\theta \middle| x\right)}\right\} = \min\left\{1, \exp\left[\frac{\varepsilon^{*2} - \varepsilon^2}{2\sigma^2}\right]\right\}, \tag{3.17}$$

where $p\left(\theta \middle| x\right)$ is the posterior distribution of the uncertain parameters, θ, given the measured production history x. The initial value of variance σ^2 is calculated based on the variance of global error of known samples (cases with actual simulation results). ε^{*2} and ε^2 are global error of the proposed uncertain parameter realization and one in current step, respectively, calculated from multiple-response-parameter proxies.

The MH algorithm will accept the proposed θ^* with probability of α^* and reject it with probability 1-α^*. In other words, the Markov chain will move to the next point if α^* is higher than the random number between 0 to 1. Otherwise, the Markov chain will remain at the current point. The MH algorithm continues until Markov chain is converged. Normally the number of steps of 10^5-10^6 gives the converged Markov chain. In addition, the acceptance ratio should be in 15-51%. The variance, σ^2, will be adjusted automatically in the workflow until the acceptance ratio is within this range.

Then, the initial 20% of Markov chain should be removed due to a burn-in period before the chain converges to the target distribution (Slotte and Smorgrav 2008). Lastly, the remaining MCMC ensemble is obtained for the later step of filtering for cases to be validated with reservoir simulator.

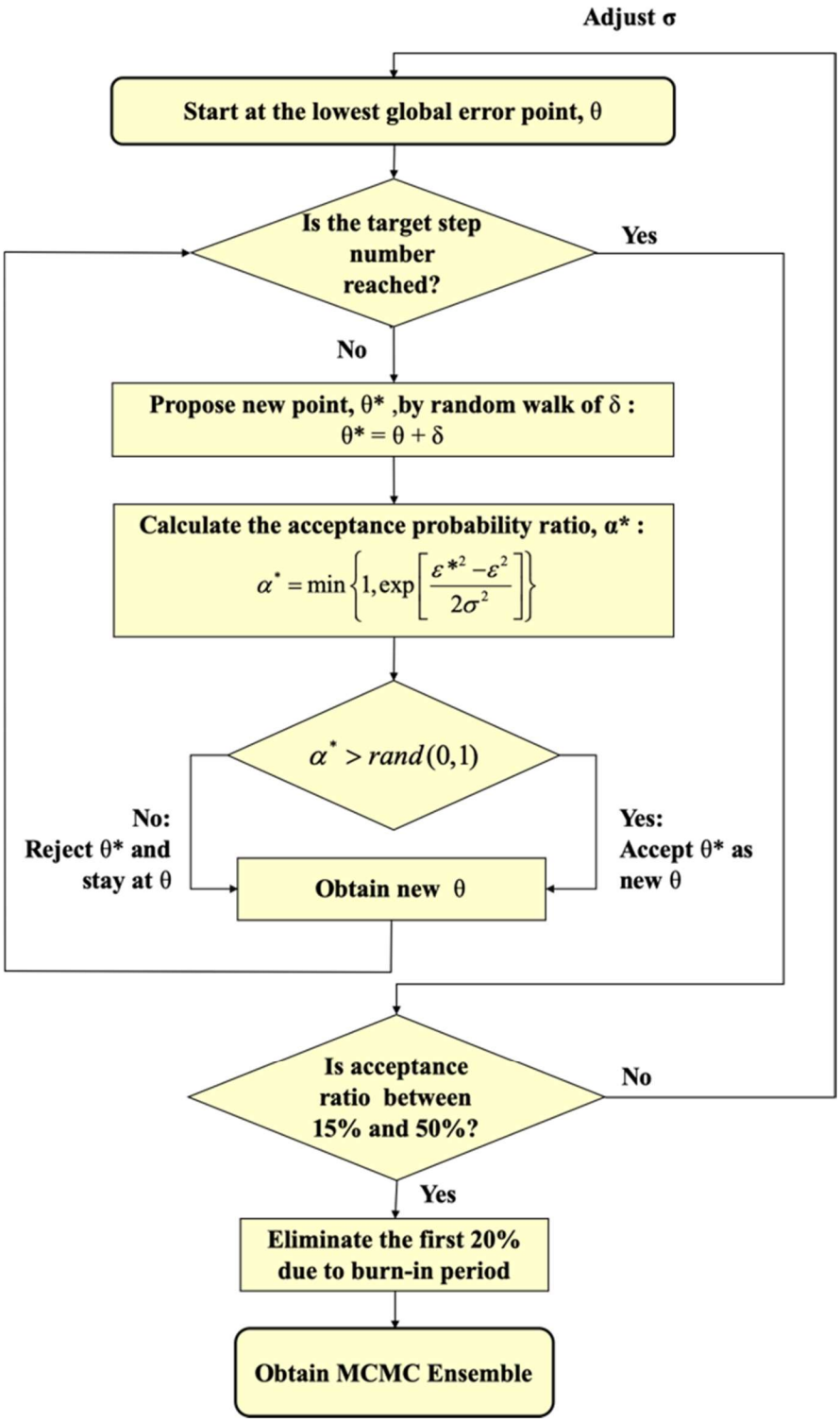


Figure 3.6: Steps and formulation in Metropolis-Hasting (MH) MCMC algorithm.

3.6.2 Proxy-Based MCMC Algorithm and Stopping Criteria

One issue with the traditional MCMC is that it requires the simulation runs for all those 10^5-10^6 steps, which are not feasible for a practical use for history matching. Therefore, proxy model is utilized together with MCMC in this study, as shown in Figure 3.7. Either K-Nearest Neighbors or Neural Network is constructed as proxy model to minimize a number of actual reservoir simulation runs. While history matching is a highly non-linear problem, constructing one proxy model to predict objective function values may be challenging to handle the nonlinearity of the problem. Therefore, multiple proxies are constructed to represent each production data point such as gas rate and water-gas ratio at different selected date. Then, these multiple points are calculated to obtain global objective function value. There are some rules of constructing multiple proxies, firstly, number of proxies should be higher than number of input parameters and secondly, constructing more proxy models should give more information to the global objective function calculation such as different production trend (Goodwin, 2015). Using Proxy-based MCMC algorithm, history matching can be performed automatically in an iterative manner, and at the end of each iteration, new proposed cases such as 25 cases are filtered from MCMC ensemble obtained from the algorithm to be validated with reservoir simulator.

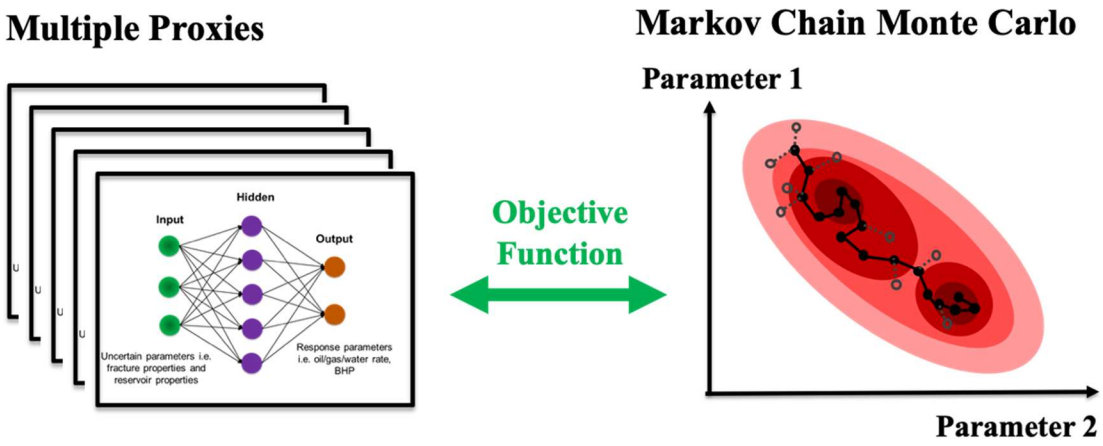


Figure 3.7: Proxy-based MCMC algorithm illustration.

The first filtering scheme for MCMC ensemble is the lowest 10% of proxy-estimated global error criteria. The reason is that more points in low global error area are needed to improve proxy quality (adding more measured points from actual simulation runs in the possible HM solution domain). Because more accuracy in the low global error domain is needed in order to better estimate the objection function value from proxy. Then, the remaining ensembles are further filtered by two criteria. For example, if the final number of cases to be validated with reservoir simulator is 25. The first half of 25 is filtered by the criteria 1 of the lowest objective function value and the latter half is filtered by the criteria 2 of the farthest points from all previous measured points. The reason we include two criteria is that first to find the HM solutions based on proxy-estimated values (criteria 1) and second to improve proxy quality by scattering the points over the uncertain-parameters domain at the same time (criteria 2). This means the workflow would find more possible HM solutions in other uncertain parameter domains. Finally, the candidate cases obtained at the end of each iteration are validated with reservoir simulator.

Then, the new actual simulation results are included into the data set for further re-training the proxy models. This algorithm continues iteratively until the stopping criteria is triggered. In this study, the maximum number of simulations and proxy convergence are used. The reason that the proxy convergence is used is because no more resources should be utilized further if proxy model improvement is minimal. For the calculation of proxy convergence, we evaluated whether the current proxy negligibly changes from the previous proxy. 100,000 random uncertain parameter realizations will be generated. Next, proxy-estimated global errors from both current and previous proxy are calculated and compared. The percentage of relative discrepancies of global error between two proxies (current and previous proxy) of all 100,000 random cases are averaged and evaluated whether it reaches the specified proxy convergence threshold. In this workflow, 5% relative discrepancy was used as the threshold.

3.7 Steps in Assisted History Matching (AHM) Workflow

There are three main steps in the AHM workflow including (1) parameters identification and screening, (2) history matching and (3) probabilistic forecasting. The main steps of AHM workflow are illustrated in Figure 3.8. At the end of the first step of parameter identification and screening, the significant parameters and objective function formulation should be determined. Then, at the end of history matching step, multiple solutions and posterior distribution of uncertain parameters can be obtained and lastly, at the end of probabilistic forecasting, P10-50-90 Estimated Ultimate Recovery (EUR) should be available based on production prediction. The detailed AHM workflow is also shown in Figure 3.9.

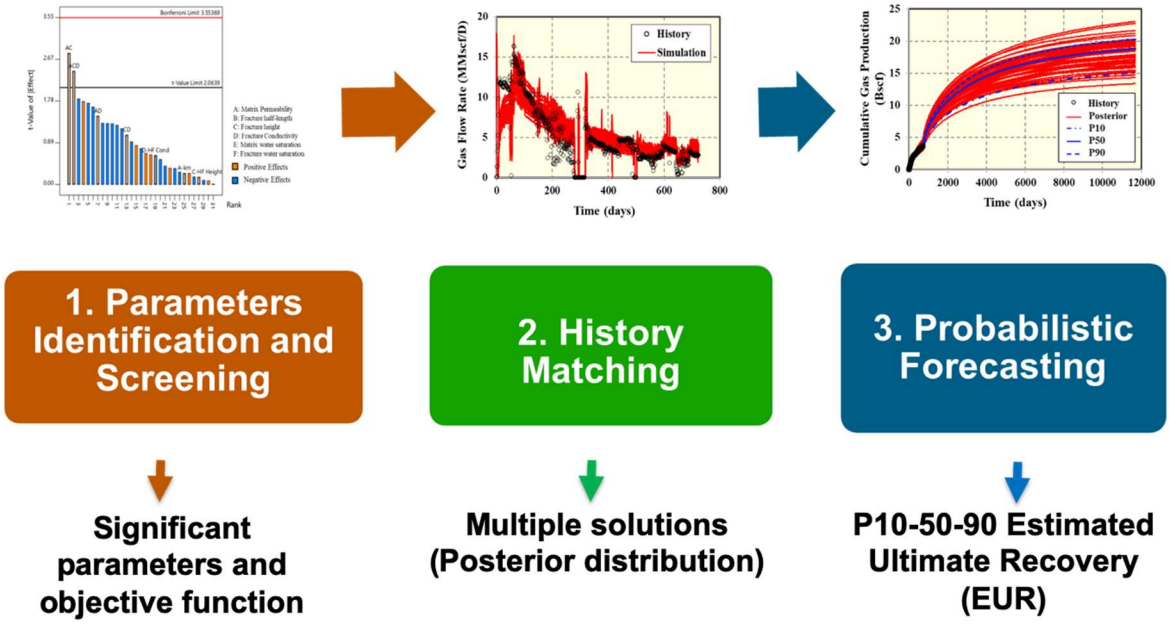


Figure 3.8: Main steps in Assisted History Matching (AHM) workflow.

3.7.1 Parameters Identification and Screening

There are two types of parameters consisting of uncertain parameters and response parameters. First, uncertain parameters should be identified. They are subsurface parameters that have uncertainty and need to be characterized by the dynamic data; for example, matrix permeability, fracture conductivity and fracture geometry. Then, response parameters are determined, which are the flowing bottomhole pressure (BHP) and historical production data used in history matching, for example: oil, gas and water flow rate.

Multiple Objective Functions

First, the objective function is defined. In this study, the multiple objective functions are used and defined either as Equation 3.18 or Equation 3.19.

Objective Function = global error (ε) =

$$\frac{\sum_{j=1}^{m}\sum_{i=1}^{n}\left|\frac{x_{ij,\mathrm{model}}-x_{ij,history}}{x_{ij,history}}\times 100\right|\times w_{ij}}{\sum_{j=1}^{m}\sum_{i=1}^{n}w_{ij}}, \qquad (3.18)$$

where i is an index of data point, j is an index of response parameter, n is a number of data points, m is a number of response parameters, $x_{ij,model}$ is the value from model (simulation run or proxy model) at index i of response parameter j, $x_{ij,history}$ is the historical data at index i of response parameter j, and w_{ij} is the weight of data points i of response parameter j.

Objective Function = global error (ε) =

$$\frac{\sum_{j=1}^{m}\sum_{i=1}^{n}\left[\frac{\left(x_{ij,\mathrm{model}}-x_{ij,history}\right)}{NF_j}\times 100\right]^2\times w_{ij}}{\sum_{j=1}^{m}\sum_{i=1}^{n}w_{ij}}, \qquad (3.19)$$

where i is an index of data point, j is an index of response parameter, n is the number of data points, m is the number of response parameters, $x_{ij,model}$ is the value from model (simulation run or proxy model) at index i of response parameter j, $x_{ij,history}$ is the historical data at index i of response parameter j, NF_j is Normalized Factor defined as the maximum of data value minus the minimum of data value of response parameter j, and w_{ij} is the weight of data points i of response parameter j.

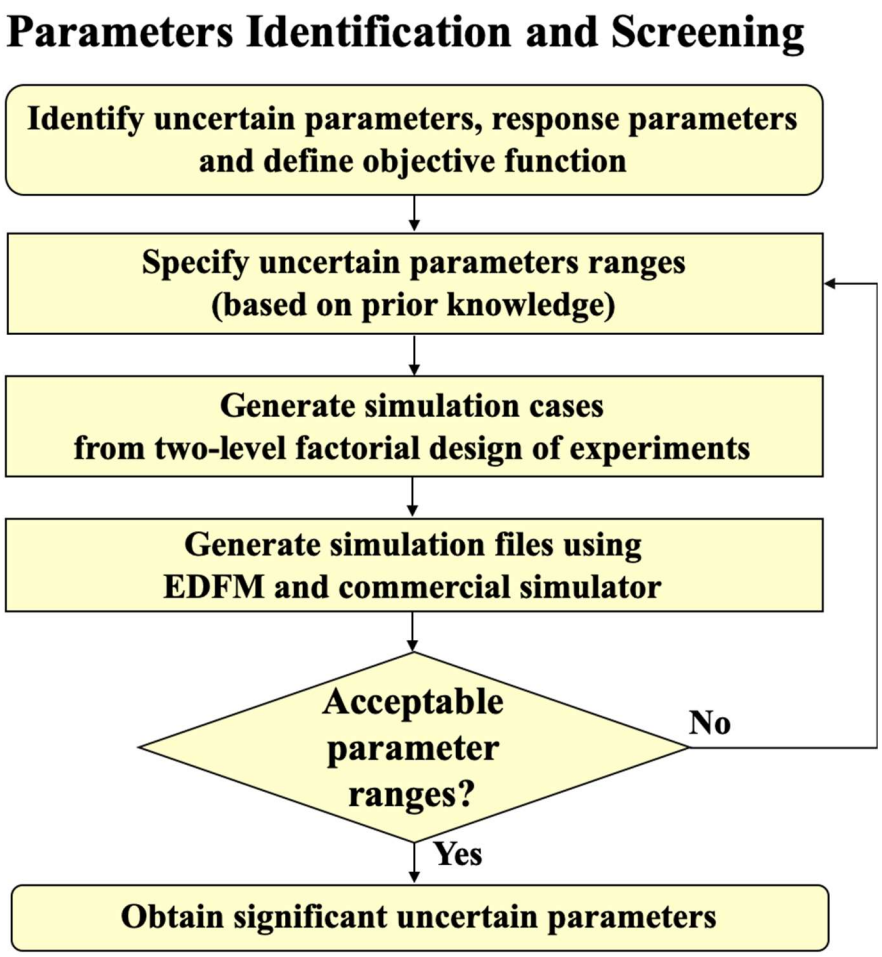
The main difference between Equation 3.18 and 3.19 is that Equation 3.19 magnifies the high relative error case to have a higher error than those of Equation 3.18 due to its squared formula. This means if we would like to distinguish the low and high error dramatically, Equation 3.19 should be used.

The concept of using global error as objective function is to have multiple objective functions into one objective function, normalized by some factor in order to eliminate the

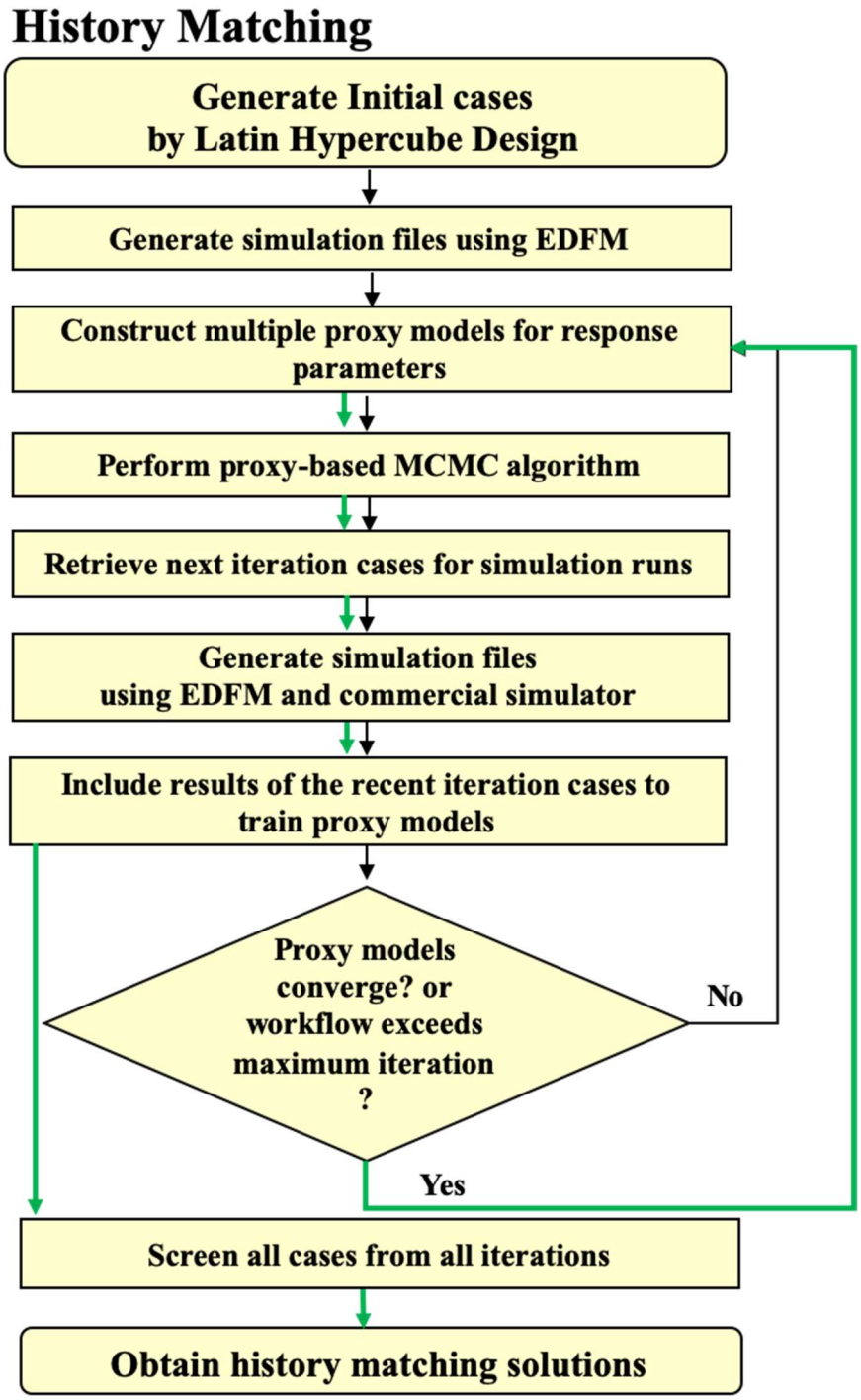
different values and ranges of each response parameter. In addition, there is a weight factor to be applied to each response parameters and further to each data point as needed (CMG-CMOST, 2017; Christie et al., 2013). In sum, the formulations can represent multiple objective functions into one value as the global error. Many response parameters such as oil flow rate and water cut at different production dates can be used.

After the objection function is defined, the effective ranges of each uncertain parameter are specified based on prior knowledge. Next, the two-level factorial design of experiments is performed in order to screen significant uncertain parameters for history matching. In this workflow, the EDFM was utilized to model fractures. After obtaining the response values of two-level factorial design, in this case global error, we evaluated whether any transformation was needed for the analysis by using Box-Cox plot (Box and Cox, 1964). Then, the significant uncertain parameters can be identified from the Pareto chart of t-value. Uncertain parameters with higher t-value will have more effect on the objective function. Ones above Bonferroni limit should be selected due to their certain importance to objective function. Ones above t-value limit are possibly important and depend on user whether to include them or not. The rest below t-value limit are the insignificant parameters and should be selected only to support hierarchy.

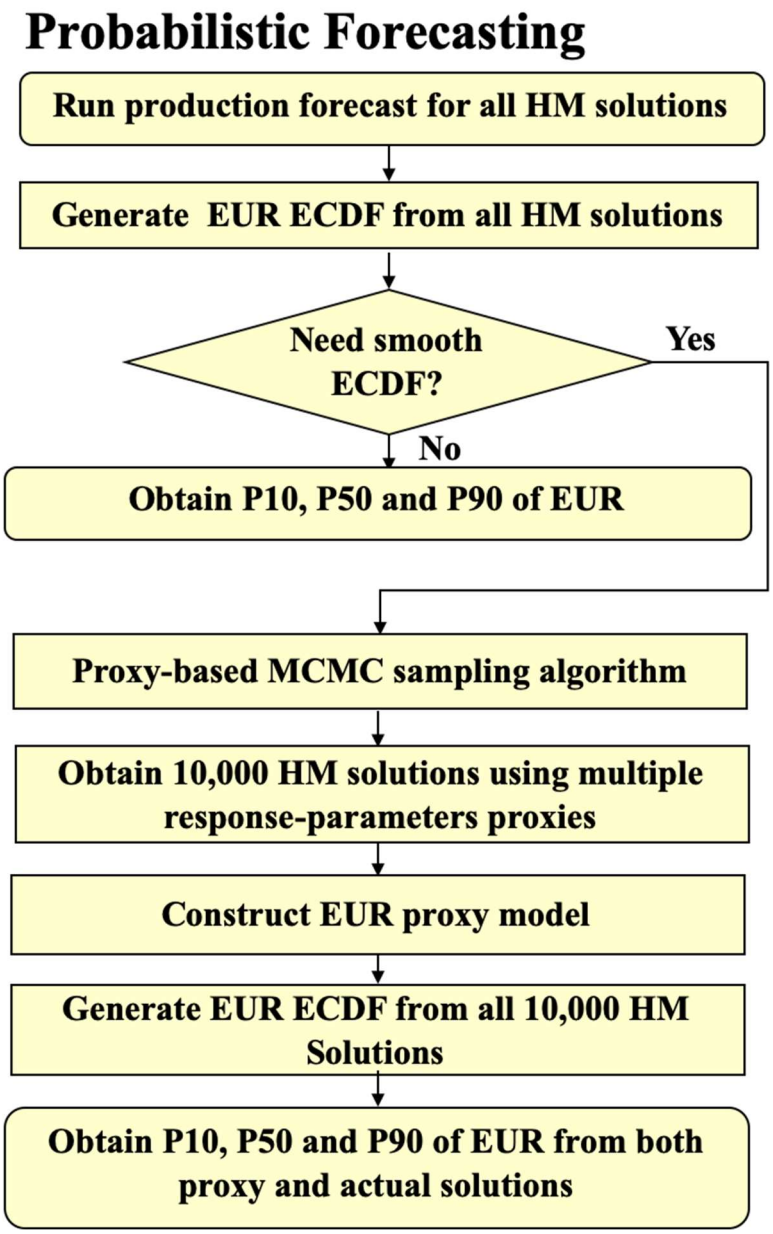
In addition, one thing to check is whether the specified parameter ranges generate the simulated production results covering all the historical data range or not. If they fail, the parameter ranges have to be adjusted and the workflow in this step should be run again. The benefits of having the parameter-screening step are that the number of simulation runs can be optimized in history matching step by eliminating the insignificant uncertain parameters.



(a) Parameters identification and screening



(b) History matching



(c) Probabilistic forecasting

Figure 3.9: Steps in AHM workflow: (a) parameters identification and screening, (b) history matching and (c) probabilistic forecasting.

3.7.2 History Matching

After identifying the significant uncertain parameters, the initial simulation cases are generated by Latin Hypercube Design. The concept of Latin Hypercube sampling (LHS) is to divide the domain into squares in case of 2 parameters or cubes for 3 parameters. The same concept is also applicable to any higher dimension. Then, the LHS evenly samples the points all over each square or cube as shown in Figure 3.10, unlike MC sampling that will sample randomly in the domain. This makes Latin Hypercube Design generate a good spatial location of initial points for proxy modeling by scattering all initial points all over parameters domain (Goodwin, 2015; Yang et al., 2015) unlike MC or the two-level factorial design that evaluates at the maximum and minimum values only. There is no strict rule for the number of initial points because it depends on the complexity of the problem and number of uncertain parameters. In this study, based on experience and empirical observation, we used number of initial points around five times of number of uncertain parameters. However, the concept is that we would not like to waste too much simulation points at the initial iteration in the domain that have high objective function values, which is the non-HM solution domain. More points should be added into the possible HM solution domain (Wantawin et al., 2017a).

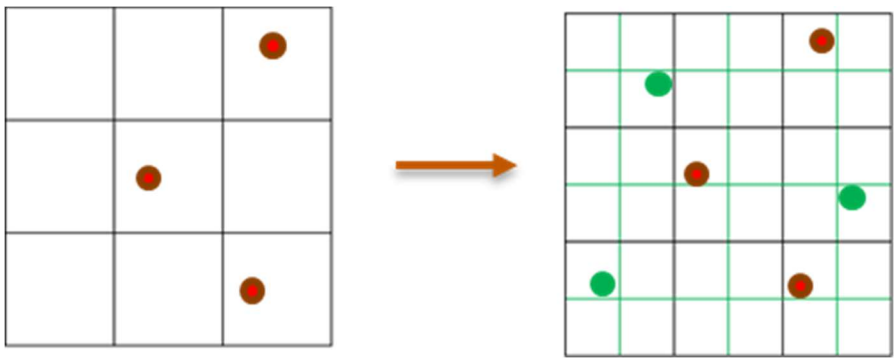


Figure 3.10: Latin Hypercube sampling (LHS) concept

Then, the multiple-response-parameter proxies will be constructed based on a training data set of the initial iteration cases. Then, proxy-based MCMC algorithm is performed automatically. As explained in detail in section 3.6.2, the algorithm continues iteratively until the stop criterion is triggered. Then, the last iteration is generated but with different filtering scheme to obtain the cases for the next iteration. That is the remaining MCMC ensemble will be screened only with the criteria 1 of lowest objective function value for the last iteration. The reason is that the final proxy is used to only obtain another set of possible history matching solutions cases and not to consider improving the proxy accuracy using the criteria 2 of the farthest points from all previous measured points as in the previous iterations. Finally, all simulation runs from all iterations are screened by HM tolerance to obtain HM solutions.

3.7.3 Probabilistic Forecasting

After multiple HM solutions are retrieved from history matching step, production forecast based on these solutions can be performed. Then, the forecasted Estimated Ultimate Recovery (EUR) of all solutions are retrieved as Empirical Cumulative Distribution Function (ECDF). The P10, P50 and P90 of EUR are quantified. In some case, the number of solutions is insufficient to create the smooth EUR ECDF. The MH MCMC algorithm can be used to generate different 10,000 solutions from the latest multiple-response-parameter proxies. This is to collect the realizations that the proxy-estimated global error is lower than HM solution tolerance until 10,000 solutions are reached. Then, another proxy for EUR is created from the actual simulation EUR retrieved previously. Next, the EUR of 10,000 solutions will be estimated from EUR proxy. Finally, a smoother ECDF and P10-P50-P90 of EUR for proxy solutions can be obtained and compared to the ECDF from the actual solutions.

Chapter 4: Application of the Assisted History Matching Workflow to Shale Gas Well[1]

4.1 INTRODUCTION

The objective of this chapter is to illustrate how the proposed Assisted History Matching (AHM) workflow can be applied to the real field case. We emphasized step-by-step of the AHM workflow including parameter identification and screening, history matching, and probabilistic production forecast. We utilized the multiple proxy-based Markov Chain Monte Carlo (MCMC) algorithm in the history matching steps with k-nearest neighbors (KNN) as proxy model and Embedded Discrete Fracture Model (EDFM) as fractures model. In this chapter, we illustrated the application of the AHM workflow to a shale gas well in China with the 724-day historical production data. We had direct measurements of production data including gas and water flow rates while the bottomhole pressure (BHP) data were estimated from the tubing head pressure with pressure-loss correlation.

4.2 RESERVOIR MODEL

The field case study is a shale gas well in China. The well is a horizontal well with a length of 5800 ft. The well was completed with 20 stages of hydraulic fractures and in each stage, three perforation clusters were shot (except for the first and last stage with 2 perforation clusters). The total number of hydraulic fractures is 58.

[1] This chapter is based on previous research performed by the author, as seen in Tripoppoom, S., Yu, W., Huang, H., Sepehrnoori, K., Song, W., and Dachanuwattana, S., 2019. A Practical and Efficient Iterative History Matching Workflow for Shale Gas Well Coupling Multiple Objective Functions, Multiple Proxy-based MCMC and EDFM. *Journal of Petroleum Science and Engineering* **176**: 594–611.

Normally, reservoir engineer receives a static reservoir model built from the integrated information from geologist, geophysicist and petrophysicist to represent the complexity of shale reservoir – high heterogeneity especially in vertical direction in terms of porosity, permeability, total organic carbon content and water saturation (Aderibigbe et al. 2016; Amin et al. 2016). Due to the lack of this information for this study, we built a homogeneous reservoir model with constant thickness for the history matching (HM) purpose. The model dimension is 6000 ft long × 2950 ft wide × fracture height. We also assumed the fracture height as model dimension in Z axis, meaning that fractures penetrate the whole reservoir thickness. The number of grid blocks is 120, 59 and 1 in x, y and z direction, respectively. Basic reservoir and fractures properties are summarized in Table 4.1. The relative permeability curves for modeling gas and water flow are shown in Figure 4.1. For gas desorption, we used Langmuir isotherm as shown in Table 4.2. The reservoir model including 58 hydraulic fractures in a horizontal well is demonstrated in Figure. 4.2.

Reservoir description	**Value**	**Unit**
Model dimension (x × y)	6000 × 2950	ft
Number of grid blocks (x × y × z)	120 × 59 × 1	-
Initial reservoir pressure	7397	psi
Reservoir temperature	200	^{o}F
Residual water saturation	20%	-
Total compressibility	3×10^{-6}	psi^{-1}
Reservoir depth	8228	ft
Matrix porosity	12%	-
Well length	5800	ft
Number of stage	20	-
Clusters per stage (18 stages)	3	-
Clusters per stage (2 stages)	2	-
Cluster spacing	100	ft

Table 4.1: Summary of basic reservoir and fracture parameters used in this study

Langmuir isotherm parameters	Value	Unit
Inverse-pressure parameter for the Langmuir isotherm model, CH_4	0.000534	1/psi
Maximum moles of adsorbed component per unit mass of rock, CH_4	0.060148	gmole of component/lb of rock
Rock density	162.318	lb/ft^3

Table 4.2: Langmuir isotherm model parameters for the gas desorption effect

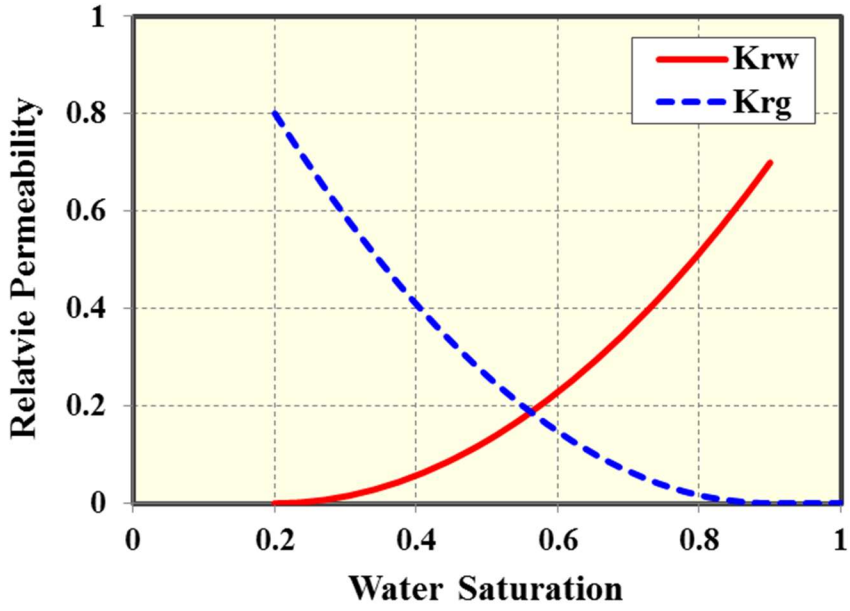


Figure 4.1: Relative permeability curves for modeling gas and water flow in shale gas reservoirs.

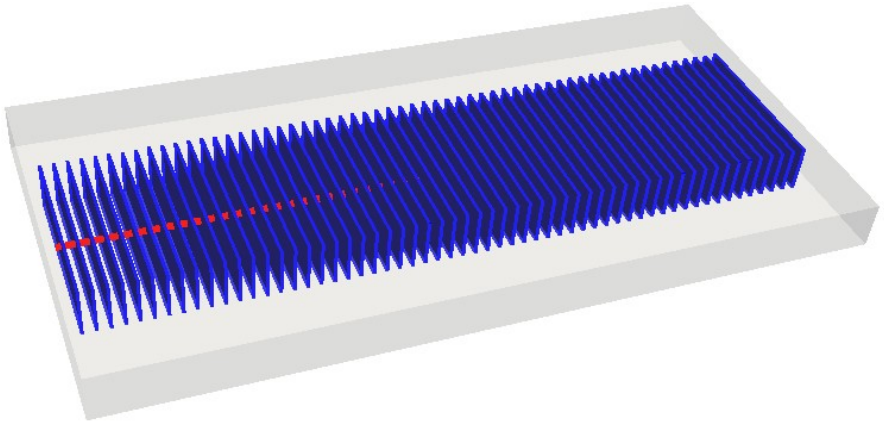


Figure 4.2: Reservoir model including 58 hydraulic fractures in a horizontal well.

4.3 Comparison between EDFM and Local Grid Refinement (LGR)

We verified the EDFM with a traditional local grid refinement (LGR) to model fractures in this field case study. We used the same reservoir and fractures properties for both methods as summarized in Tables 4.1-4.2 and other fracture and matrix parameters summarized in Table 4.3. The only difference is fracture modeling technique. For LGR, we refined the parental cells containing fractures into 7 cells in I-direction. For EDFM, we used EDFM preprocessor to calculate non-neighboring connections (NNCs) and modified reservoir simulator data file. Figure 4.3 shows the production profile results from both methods. The cumulative gas and water production for LGR are 5.35 Bscf and 15.74 MSTB while the cumulative gas and water production for EDFM are 5.47 Bscf and 15.71 MSTB. The relative differences are 2% and 0.2%, respectively. We can see that there is no significant difference between EDFM and LGR. Table 4.4 shows the CPU computational time of LGR and EDFM. EDFM can complete the run twice faster than LGR for this field case study. To conclude, EDFM can mimic the fractures flow behavior as LGR with less

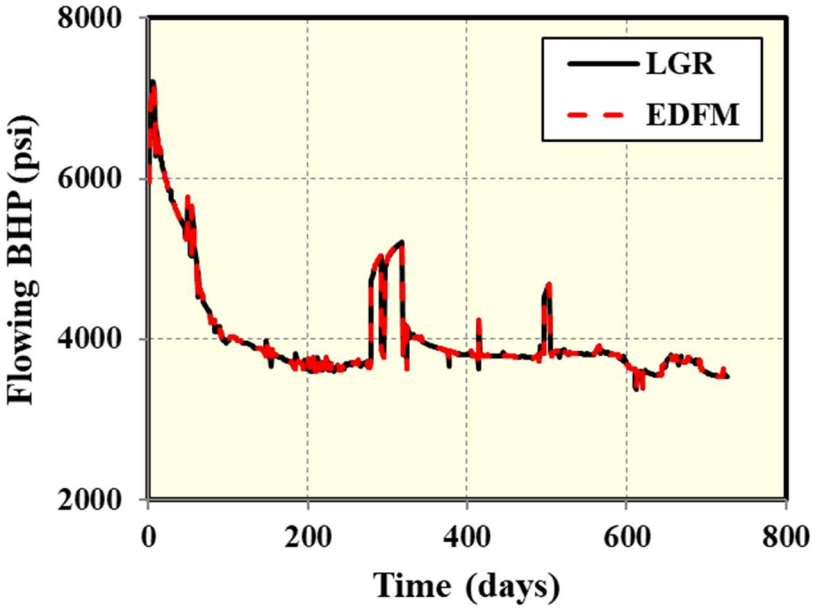
computational time for this well application. That means we can perform AHM faster for unconventional reservoirs by using EDFM.

Matrix and fractures parameters	**Value**	**Unit**
Matrix permeability	0.02005	md
Fracture width	0.01	ft
Fracture conductivity	5.5	md-ft
Fracture half-length	800	ft
Fracture height	65.5	ft
Fracture dip angle	90	degree
Initial water saturation	40%	-

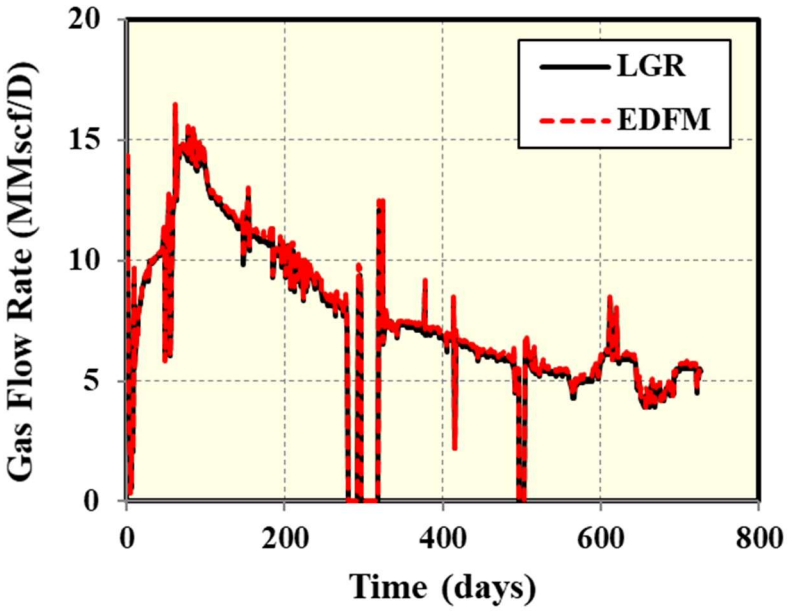
Table 4.3: Summary of matrix and fracture parameters used in EDFM and LGR comparison.

Method	**CPU time (sec)**	**Elapsed time (sec)**	**Ratio**
LGR	2318	340	1: 0.5
EDFM	1159	176	

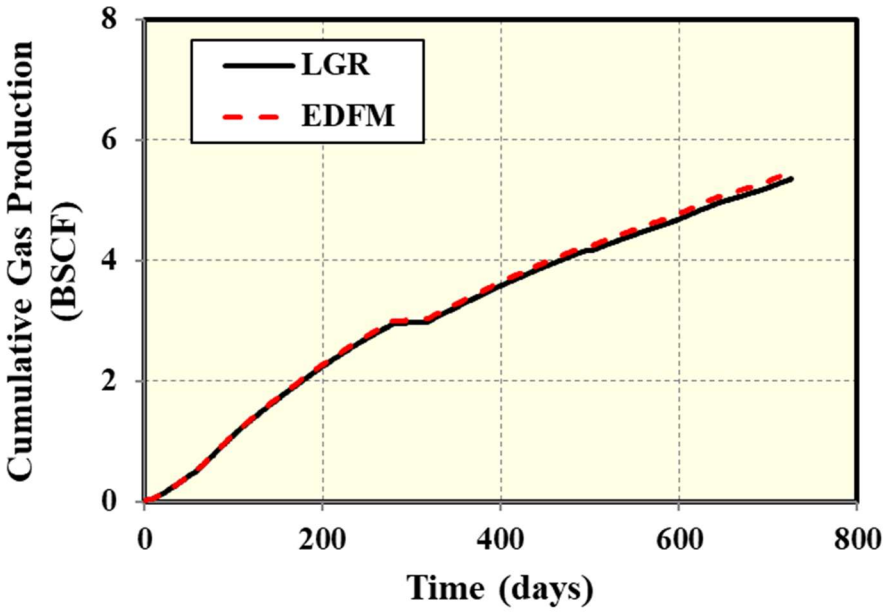
Table 4.4: Comparison of CPU time and elapsed time between EDFM and LGR



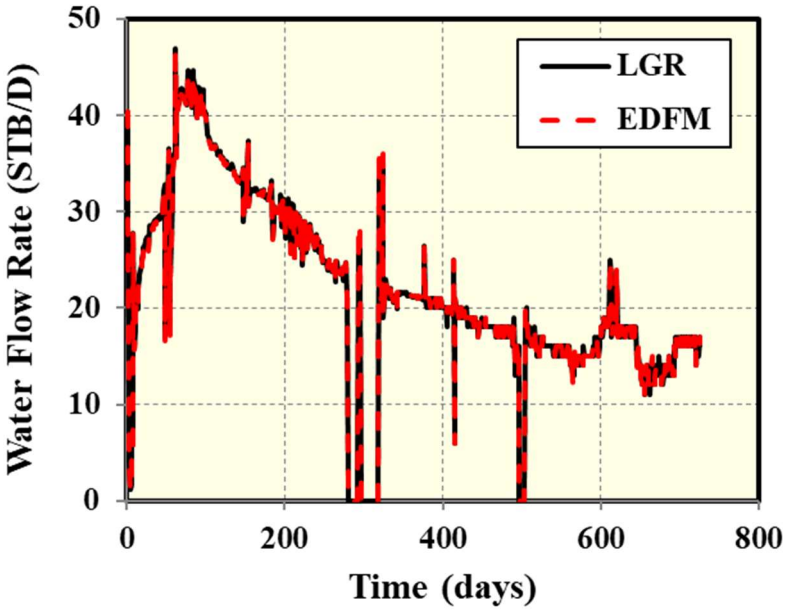
(a) Flowing BHP as constraint



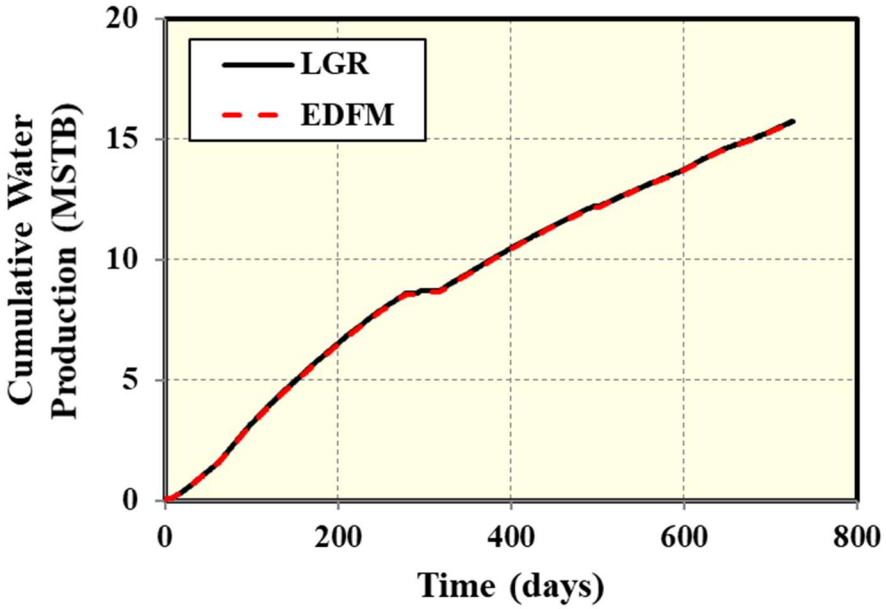
(b) Gas flow rate



(c) Cumulative gas production



(d) Water flow rate



(e) Cumulative water production

Figure 4.3: Comparison between simulation results from LGR and EDFM methods: (a) Flowing BHP as constraint, (b) Gas flow rate, (c) Cumulative gas production, (d) Water flow rate and (e) Cumulative water production.

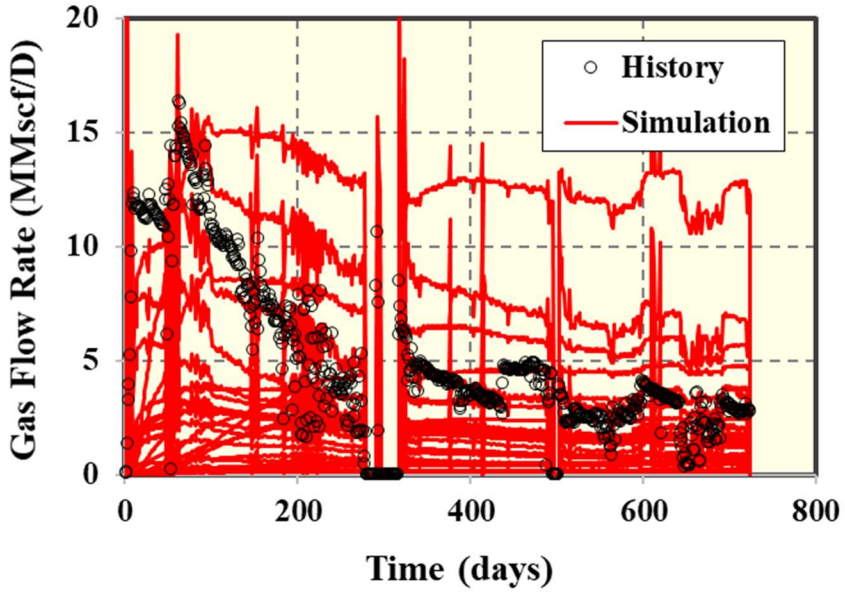
4.4 Parameters Identification and Screening

For the proposed workflow, we can identify an unlimited number of uncertain parameters for the HM purpose. It depends on how many available data we already had and what is the uncertainty of each data. In this study, we identified 6 uncertain parameters and their ranges were summarized in Table 4.5 with uniform distribution based on prior knowledge about this well. We modeled the fracture width of 1 ft and adjusted fracture permeability to have the product of fracture width and fracture permeability to be equal to fracture conductivity (Rubin 2010). We then defined the objective function as the global error of both gas rate and water rate as Equation 3.19.

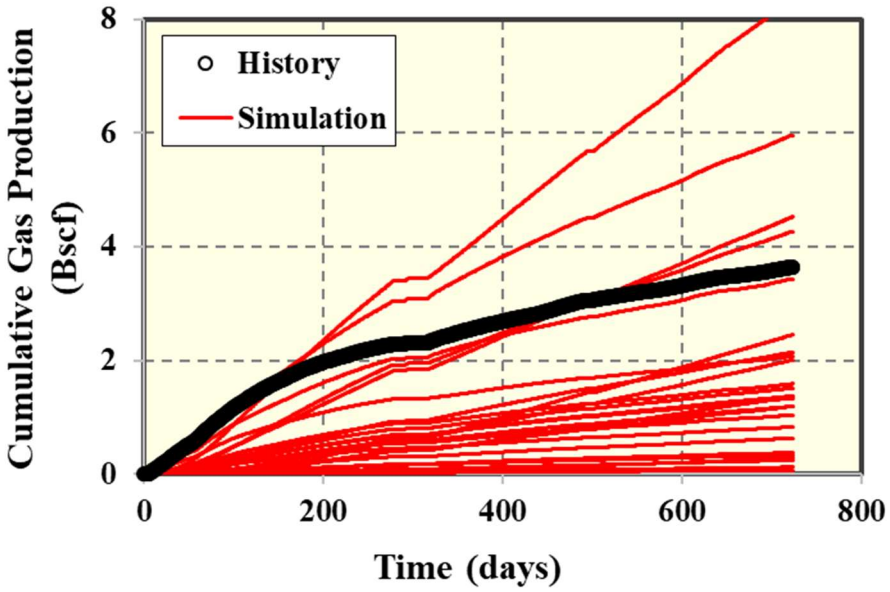
Code	Uncertain parameter	Unit	Distribution	Min value	Max value
A	Matrix permeability	mD	Uniform	0.0001	0.04
B	Fracture half-length	ft	Uniform	300	1300
C	Fracture height	ft	Uniform	16	115
D	Fracture conductivity	md-ft	Uniform	1	10
E	Matrix water saturation	-	Uniform	0.30	0.50
F	Fracture water saturation	-	Uniform	0.40	0.90

Table 4.5: Summary of six uncertain parameters and their prior distributions

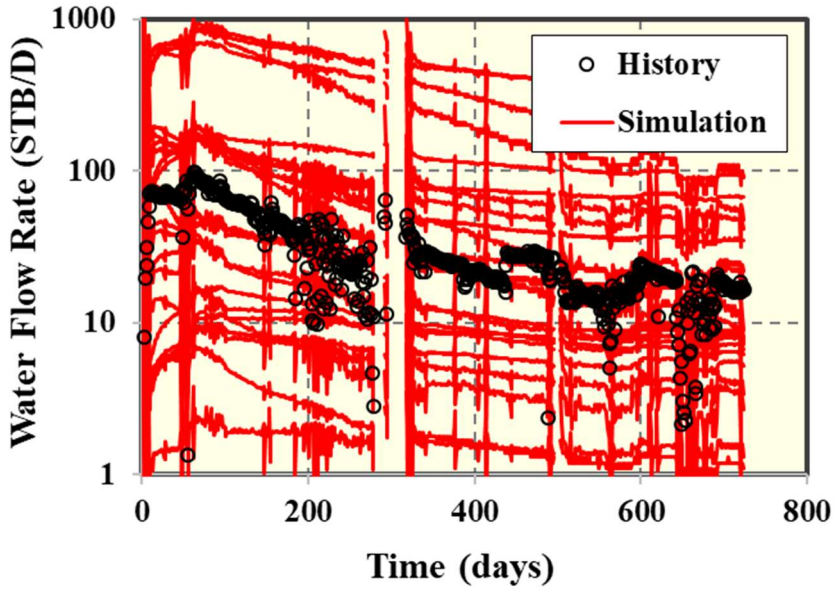
Then, we screened uncertain parameters for HM by performing two-level factorial design. Normally, the 2^6 or 64 cases should be generated for 6 parameters using two-level full factorial design. The full factorial design can analyze all parameters interactions, but it requires too many runs for the screening purpose. The 2^4 - 2^5 is sufficient for the screening step and the analysis of partial parameter interaction (Stat-Ease Inc 2017). Therefore, we used 2^5 or 32 cases to screen significant uncertain parameters. We constrained the production by BHP and monitored global error of gas rate and water rate. We used EDFM preprocessing code coupling with commercial simulator to create simulation files automatically. The results of 32 cases are shown in Figure 4.4. The simulated data covering the historical data confirms the specified ranges of 6 uncertain parameters are acceptable and do not need the revision of uncertain-parameter range at this step.



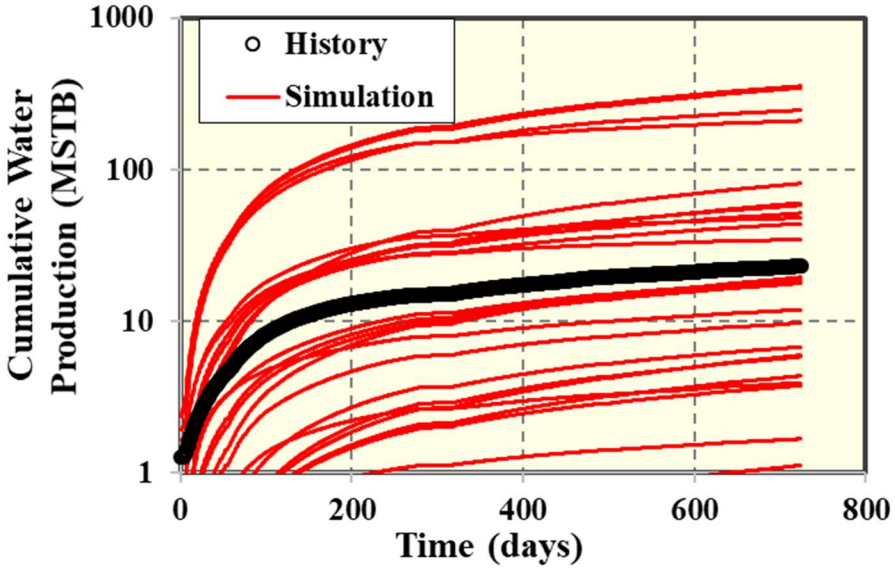
(a) Gas flow rate



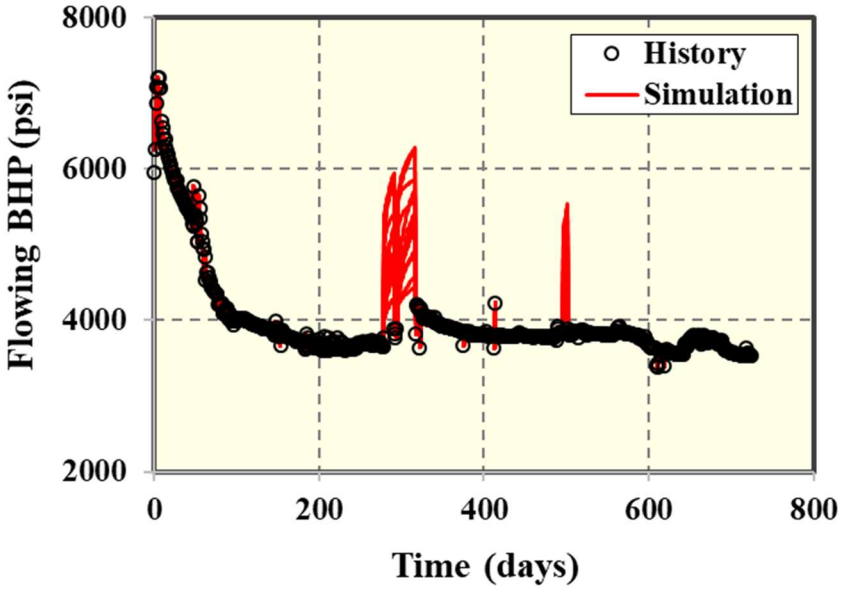
(b) Cumulative gas production



(c) Water flow rate



(d) Cumulative water production



(e) Flowing BHP

Figure 4.4: Comparison between simulation results and historical data of cases from two-factorial design: (a) Gas flow rate, (b) Cumulative gas production, (c) Water flow rate, (d) Cumulative water production and (e) Flowing BHP.

To select significant uncertain parameters, first, we evaluated whether any transformation of response values (global error) is needed by looking at Box-Cox plot (Box and Cox 1964). The lowest lambda value indicates what transformation is needed ($\lambda = 1$ for no transformation, $\lambda = 0.5$ for square root, $\lambda = 0$ for natural log, $\lambda = -0.5$ for inverse square root, $\lambda = -1$ for inverse). In this case, log transformation was most appropriate for both gas and water global error.

Then, we analyzed the Pareto chart of t-value, as shown in Figures 4.5-4.6. The Pareto chart indicated all parameters are lower than Bonferroni limit but only AC and ACD are above t-value limit, which are matrix permeability, fracture height and fracture conductivity. This means fracture parameters except fracture half-length have more impact on gas rate. For water global error, in Pareto chart, CF, DF, F, C and D are above Bonferroni

limit. The C, D and F are fracture height, fracture conductivity and fracture water saturation, respectively. Even the next one is above the t-value limit but we did not select it as significant parameter because there is a large gap between D and the next one. In sum, we screened from 6 uncertain parameters to 4 significant uncertain parameters including matrix permeability, fracture height, fracture conductivity and fracture water saturation. However, for the later demonstration in this study, we included matrix water saturation for HM.

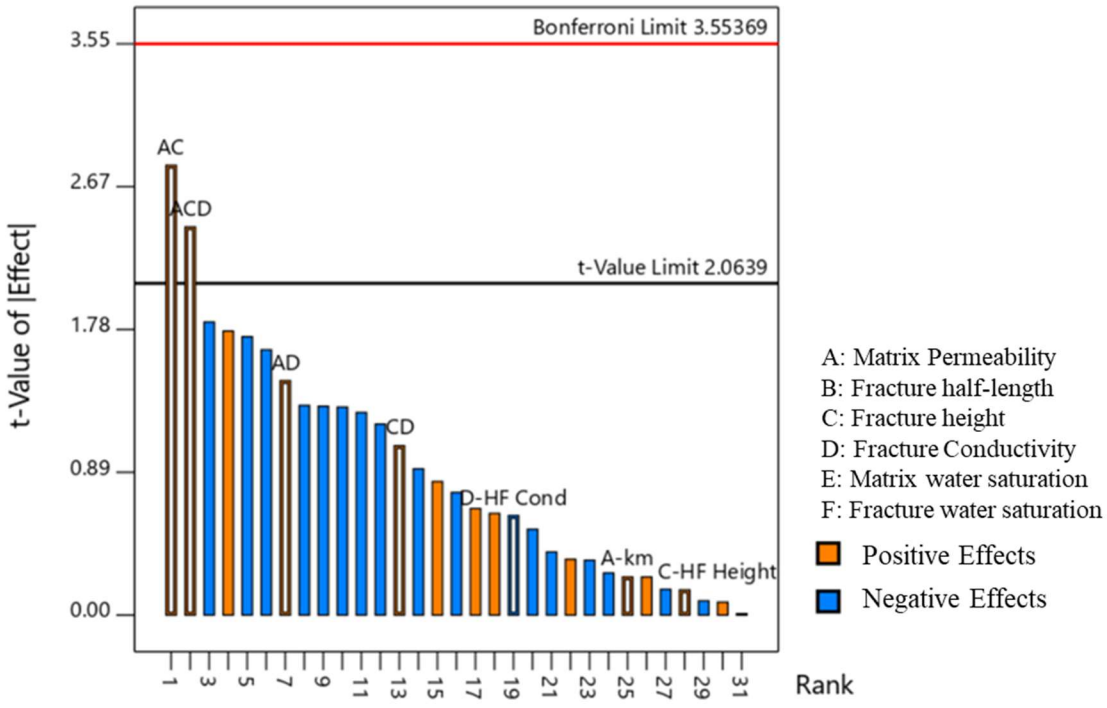


Figure 4.5： T-value ranking of two-level factorial design for global error of gas rate.

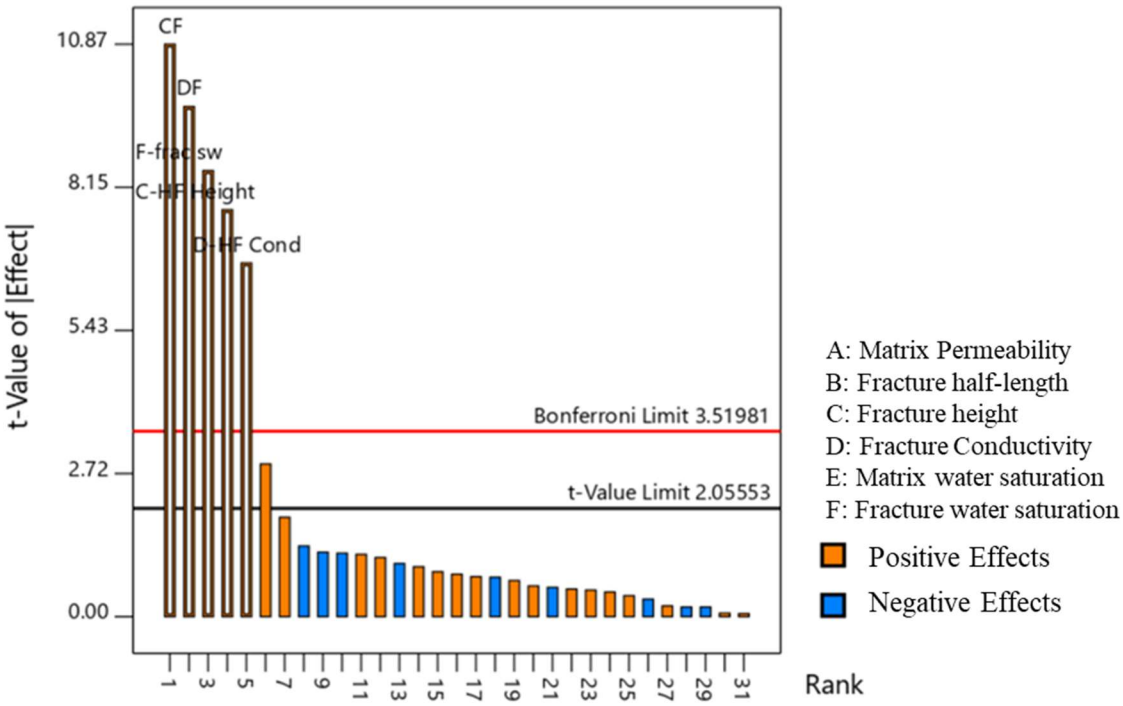


Figure 4.6: T-value ranking of two-level factorial design for global error of water rate.

4.5 History Matching

After screening uncertain parameters, we performed Latin Hypercube Design to sample the initial 25 cases from 5 uncertain parameters summarized in Table 4.6, scattering over the uncertain-parameter domain. The fracture half-length is a constant value of 700 ft for all cases in history matching. Figure 4.7 shows the results of initial 25 cases, which also cover the historical production data. Then, we created multiple response-parameter proxies from these initial 25 cases by using KNN. We created total 20 proxies including 10 proxies for gas rate and 10 proxies for water rate at days 20, 67, 199, 258, 325, 400, 508, 602, 680 and 723. Each proxy represents the different trend of historical production data. These multiple proxies will be later used in proxy-based MCMC algorithm to propose new possible realizations as HM solutions in the next iteration.

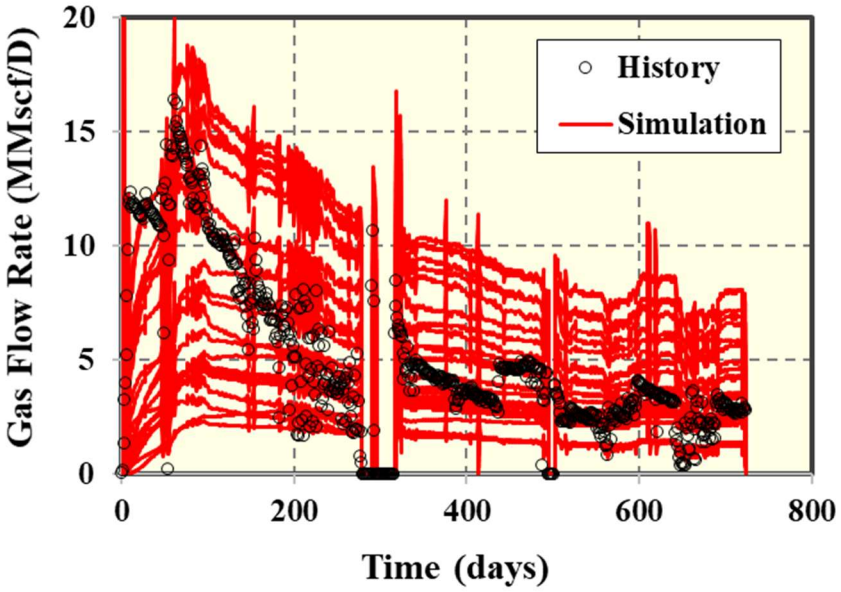
Code	Uncertain parameter	Unit	Distribution	Min value	Max value
A	Matrix permeability	mD	Uniform	0.0001	0.04
B	Fracture half-length	ft	Deterministic value	700	
C	Fracture height	ft	Uniform	16	115
D	Fracture conductivity	md-ft	Uniform	1	10
E	Matrix water saturation	-	Uniform	0.30	0.50
F	Fracture water saturation	-	Uniform	0.40	0.90

Table 4.6: Summary of the five uncertain parameters after screening step with a given fracture half-length of 700 ft

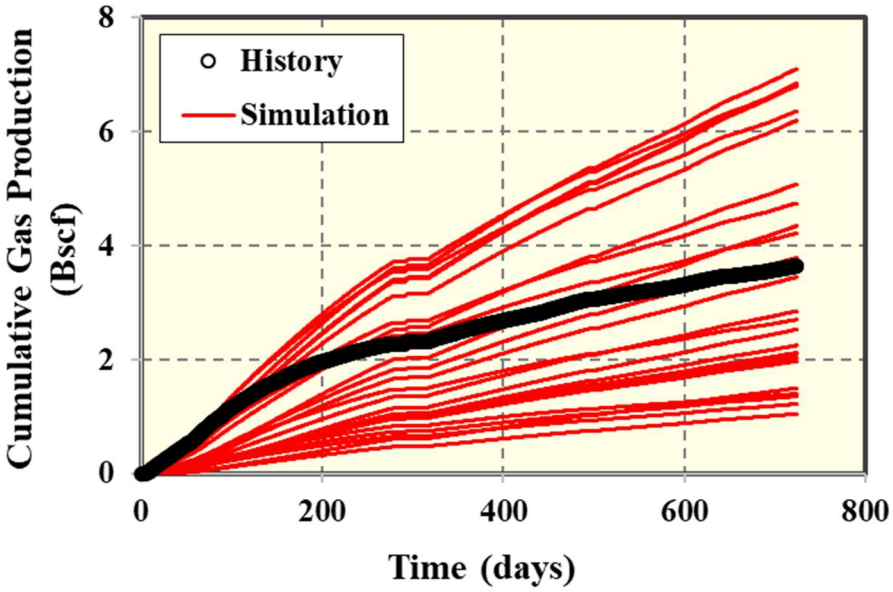
Then, the iterative HM workflow continued automatically by our developed code. Using proxy-based MCMC algorithm and filtering scheme selected the next iteration cases. We used one chain of MCMC and started the chain at the lowest global error points at each iteration. The iteration progressed until it reached the stopping criteria, which is the percentage of relative discrepancies between current and previous proxy with a threshold of 5% or the maximum iteration number (12 iterations or 300 simulation runs). For this study, the workflow stopped at iteration 11 because the percentage of relative discrepancies between current and previous proxy is below the threshold of 5%, as shown in Figure 4.8. The total computational time for each iteration - including proxy-based MCMC algorithm, simulation-running time - is around 2 hours on 4 processors, 3.0 GHz Dell computers. The total computational time for HM section is around one day.

After that, the proxy-based MCMC algorithm generated the last iteration cases with different filtering scheme using the lowest global error value for all 25 cases. Then, we evaluated all 300 cases from iteration 1 to the last iteration and screened them for HM

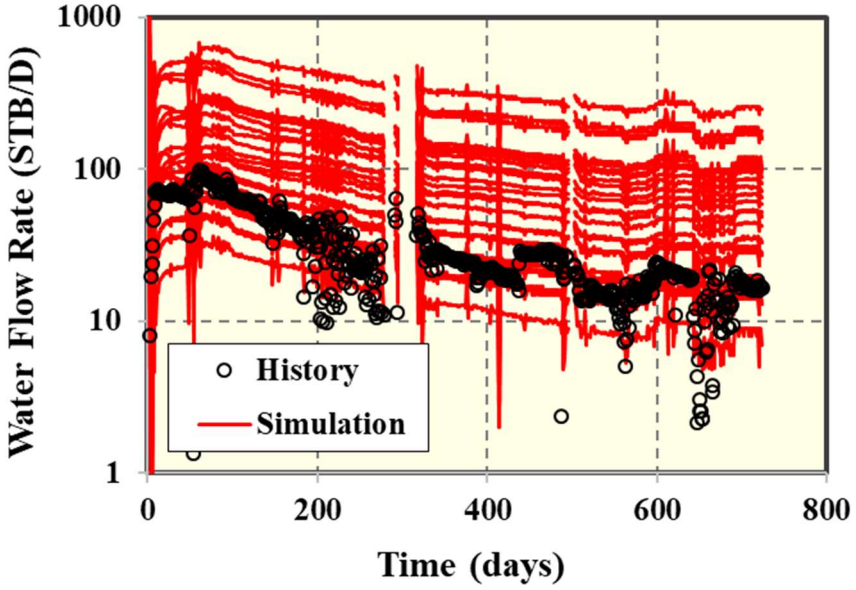
solutions. Firstly, the global error used as the threshold value for screening solutions was unknown. It can be specified only after we have the simulation results available for a comparison to production history so that we can identify which case and what maximum global error are the least acceptable as HM solutions. We used the gas global error of 220 and water global error of 220 as HM solutions threshold. 64 cases passed the threshold out of 300 simulation runs, which is approximately 21%. With more relaxed threshold, the number of HM solutions can be as high as 98 cases, which is approximately 33% out of 300 simulation runs. For this study, we used 64 HM solutions for production forecast.



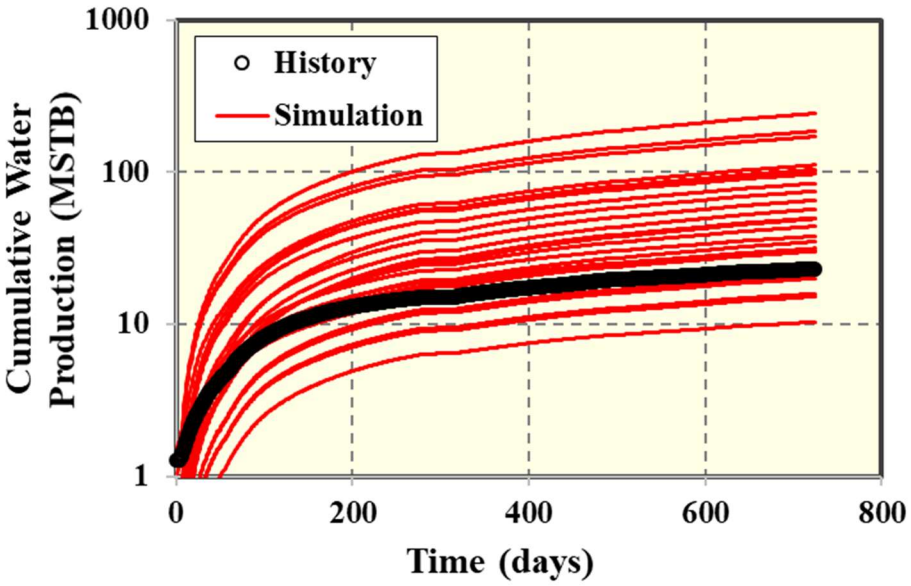
(a) Gas flow rate



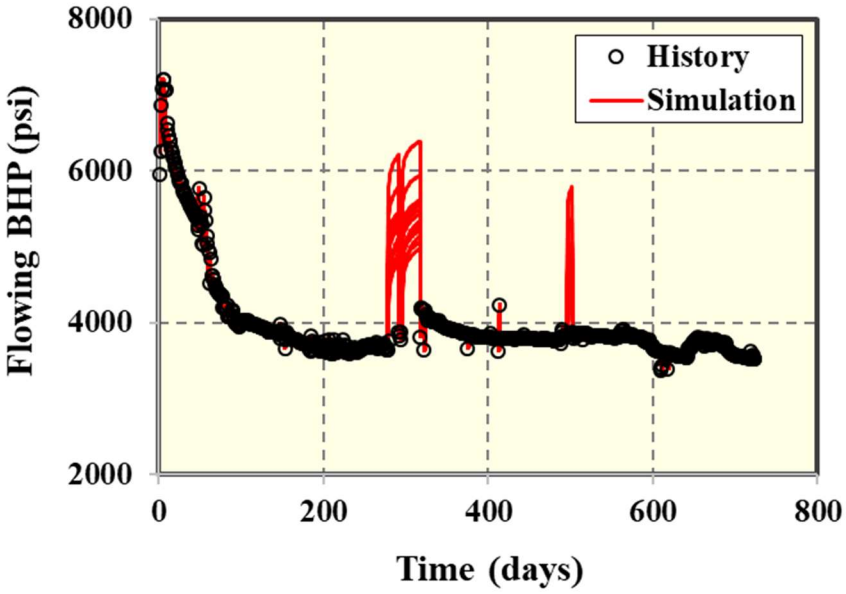
(b) Cumulative gas production



(c) Water flow rate



(d) Cumulative water production



(e) Flowing BHP

Figure 4.7: Comparison between simulation results and historical data of cases from Latin Hypercube Design (Iteration 1): (a) Gas flow rate, (b) Cumulative gas production, (c) Water flow rate, (d) Cumulative water production and (e) Flowing BHP.

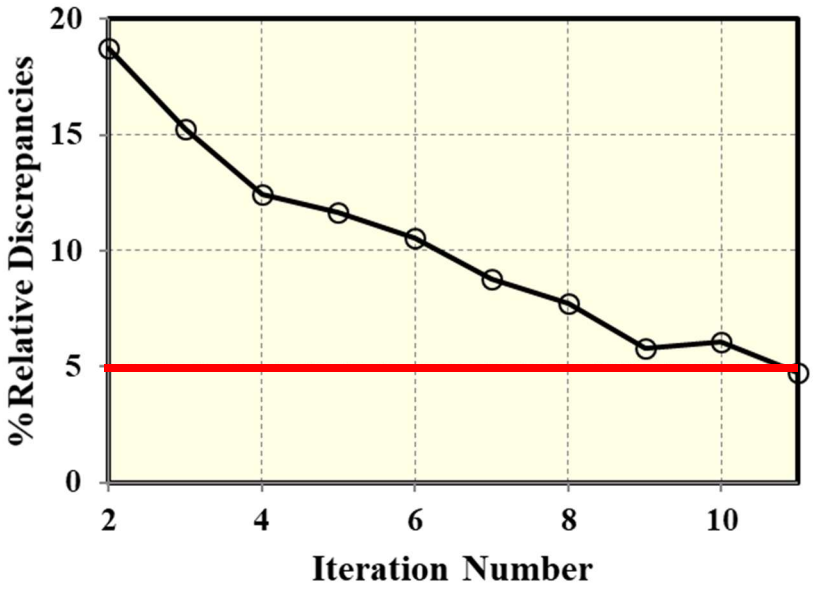
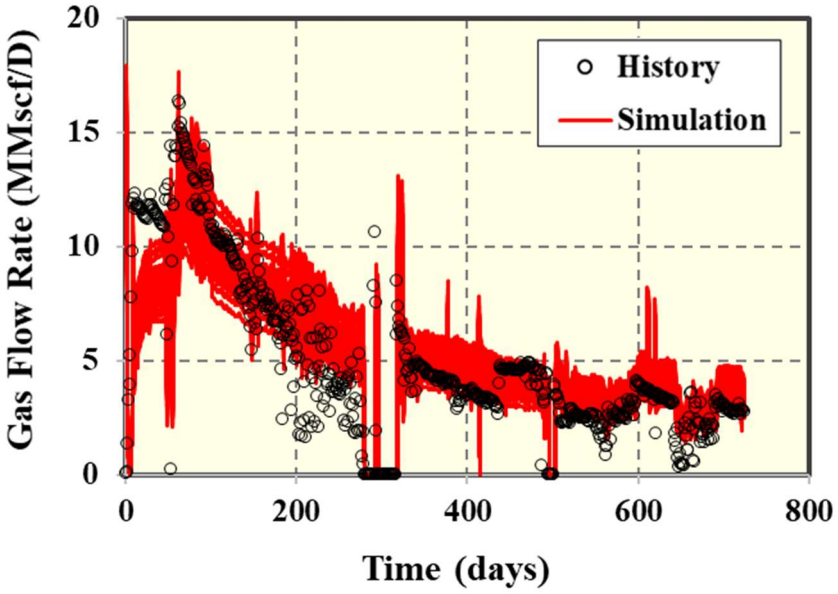


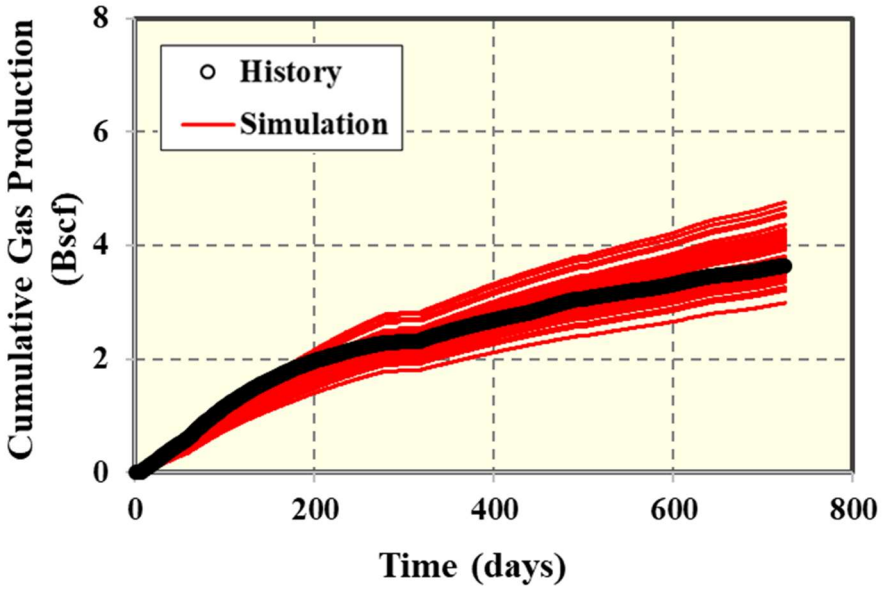
Figure 4.8: The percentage of relative discrepancies between current and previous proxy vs. iteration number.

The simulation results of 64 HM solutions are shown in Figure 4.9. The best match case are also illustrated in Figure 4.10. The best match realization has properties as follows: matrix permeability of 0.027 md, fracture height of 40 ft, fracture conductivity of 9.9 md-ft, matrix water saturation of 0.38 and fracture water saturation of 0.48. In addition, parallel plot of 64 HM solutions are also shown together with non-HM solutions cases in Figure 4.11. The prior and posterior distributions of each uncertain parameters from 64 HM solutions are shown in Figure 4.12. To get smoother posterior distribution, we also performed MCMC sampling to obtain 10000 HM solutions using multiple-response-parameter proxies. The prior and posterior distributions of each uncertain parameter from 10000 proxy-generated HM solutions are shown in Figure 4.13. Both Figure 4.12 and Figure 4.13 show a similar distribution shape. Besides, we observed the obvious change from prior uniform distribution to posterior distribution for all 5 parameters. For fracture

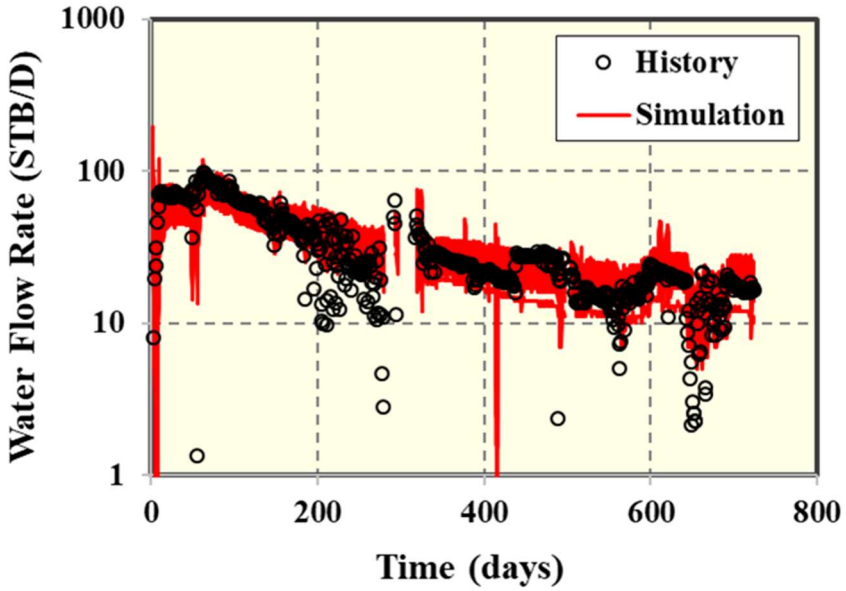
height, fracture conductivity and fracture water saturation, the posterior distribution changed significantly and had the narrower ranges. This implied that these three parameters are very sensitive to the objective function. On the contrary, matrix permeability and matrix water saturation had relatively wider ranges. This indicated that matrix permeability and matrix water saturation are not very sensitive to the objective function (Yang et al. 2015). The finding here is aligned with the parameter screening step which emphasized that matrix water saturation was not a significant uncertain parameter. The values with the highest frequency of each parameters are as follows: matrix permeability of 0.031 md, fracture height of 39 ft, fracture conductivity of 9.5 md-ft, matrix water saturation of 0.37 and fracture water saturation of 0.46.



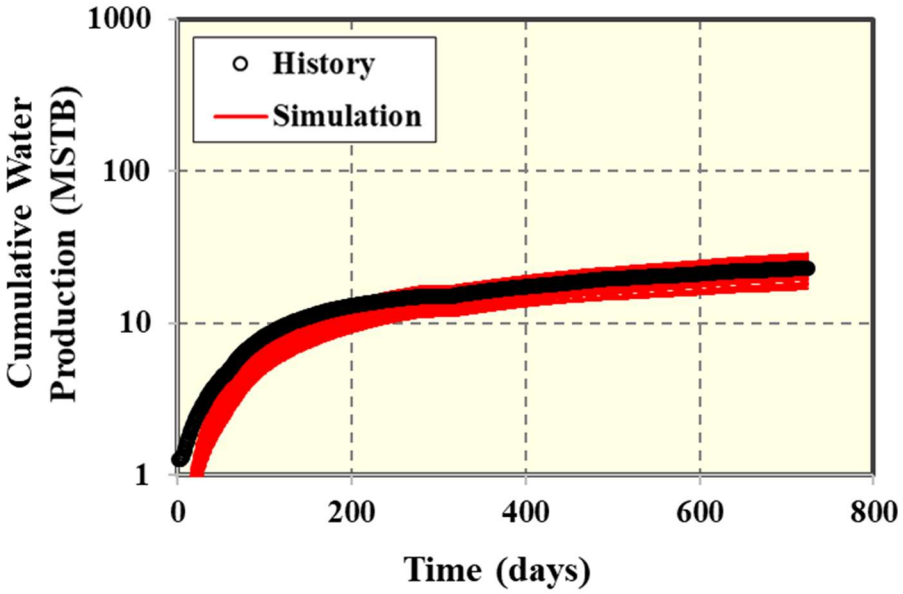
(a) Gas flow rate



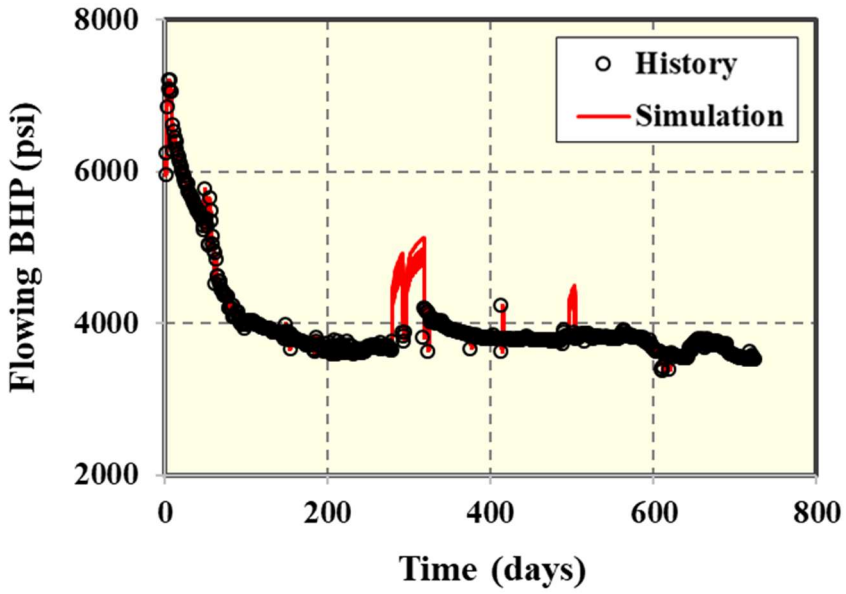
(b) Cumulative gas production



(c) Water flow rate

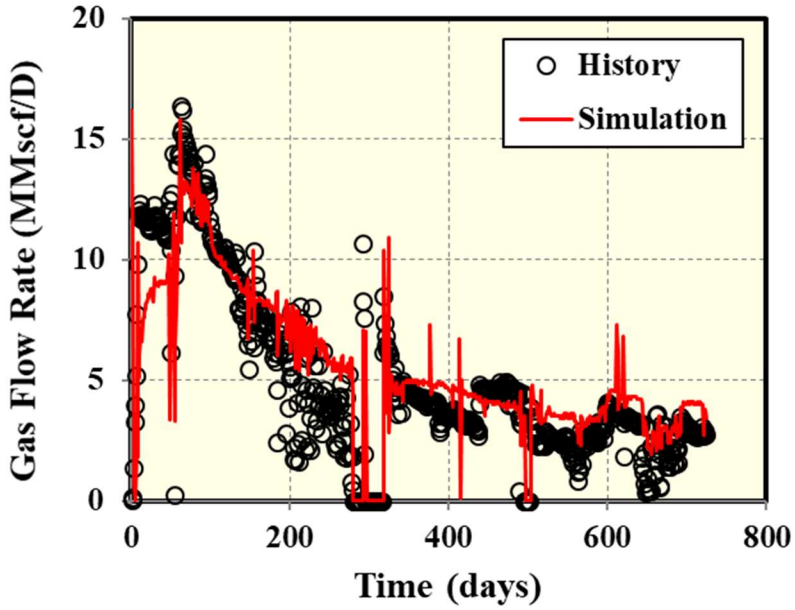


(d) Cumulative water production

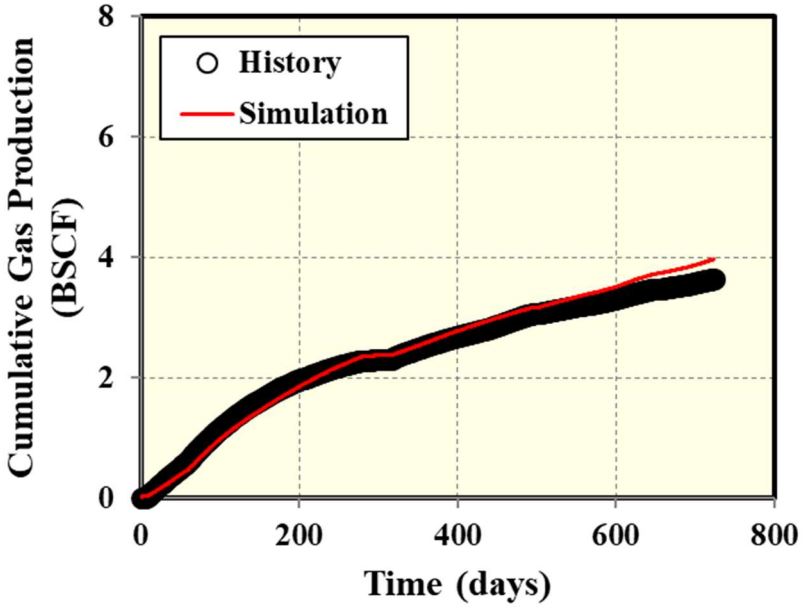


(e) Flowing BHP

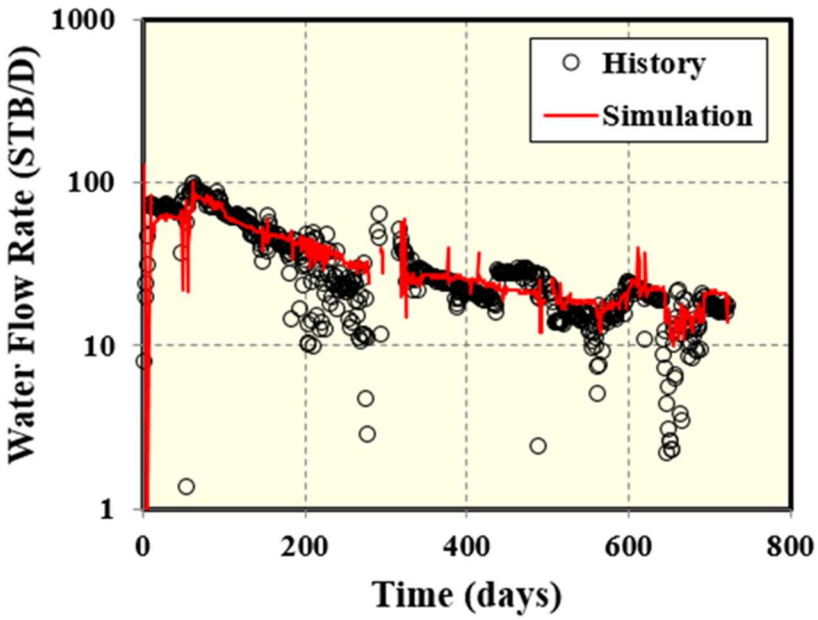
Figure 4.9: Comparison between simulation results and historical data of 64 HM solutions: (a) Gas flow rate, (b) Cumulative gas production, (c) Water flow rate, (d) Cumulative water production and (e) Flowing BHP.



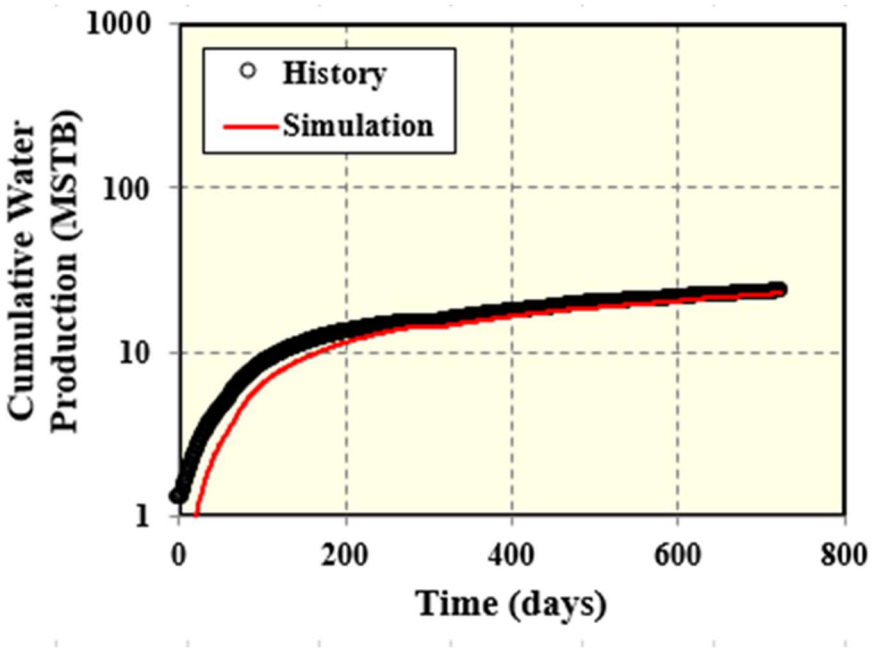
(a) Gas flow rate



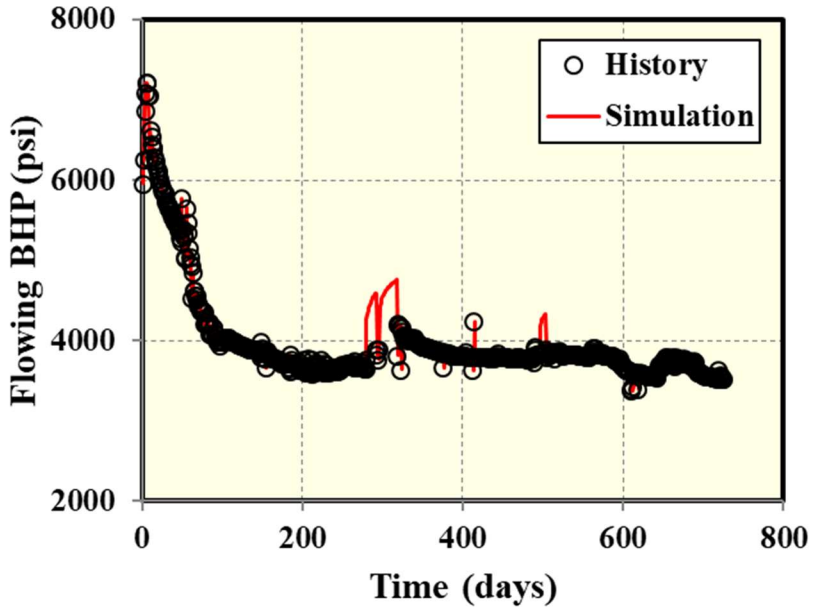
(b) Cumulative gas production



(c) Water flow rate



(d) Cumulative water production



(e) Flowing BHP

Figure 4.10: Comparison between simulation results and historical data of the best match case: (a) Gas flow rate, (b) Cumulative gas production, (c) Water flow rate, (d) Cumulative water production and (e) Flowing BHP.

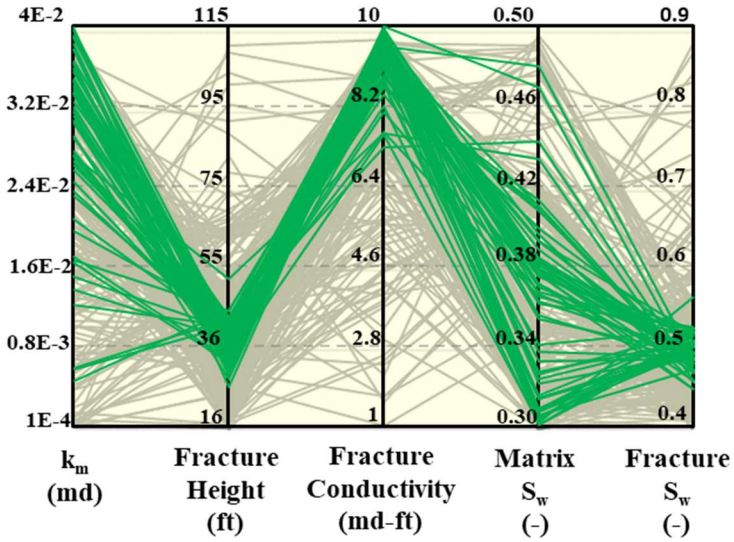
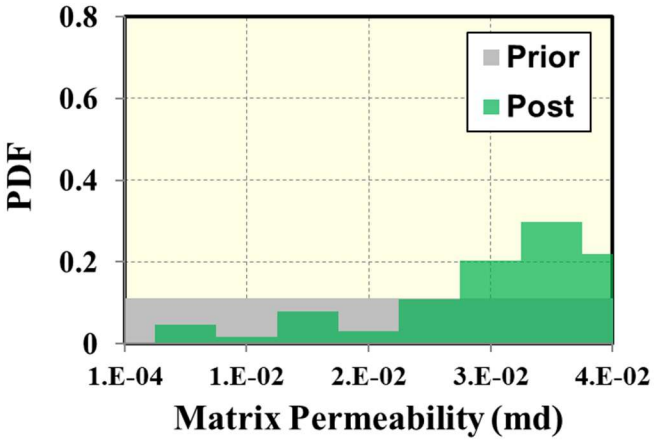
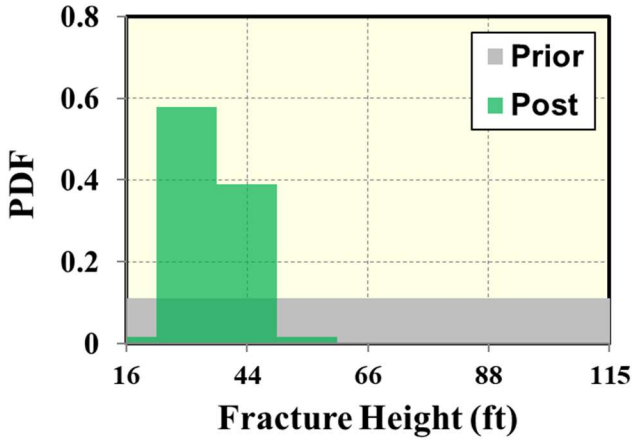


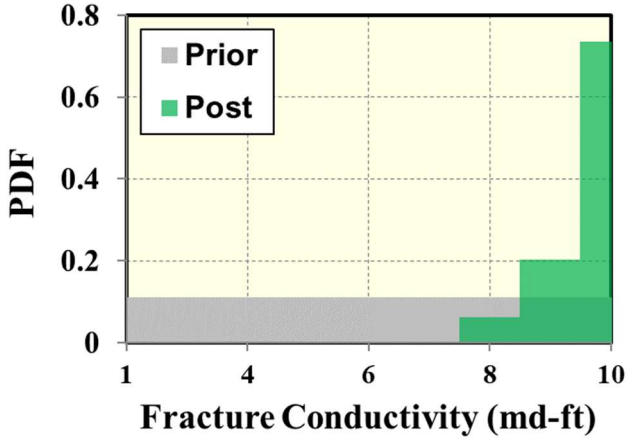
Figure 4.11: Parallel coordinate plot of 64 HM solutions. Each green line represents one realization of the solution. The grey lines represent non-HM solution realizations.



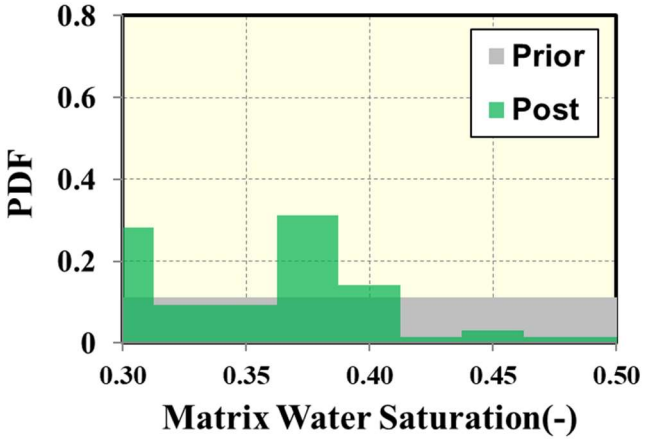
(a) Matrix permeability



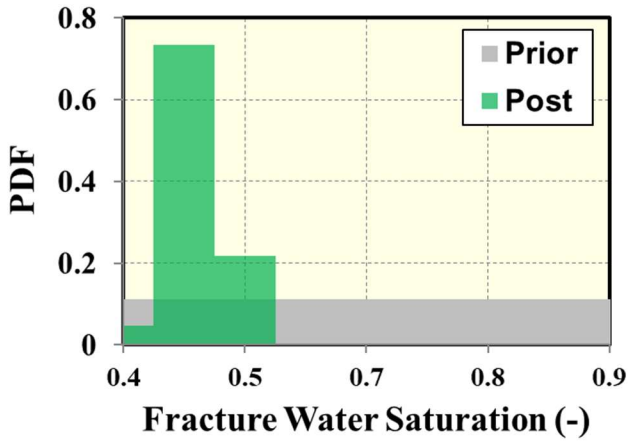
(b) Fracture height



(c) Fracture conductivity

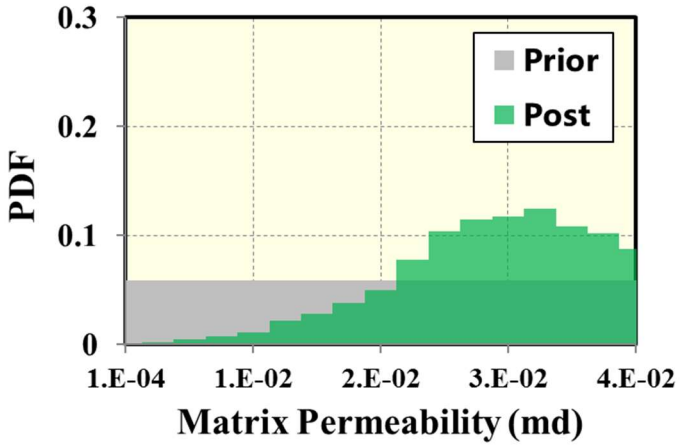


(d) Matrix water saturation

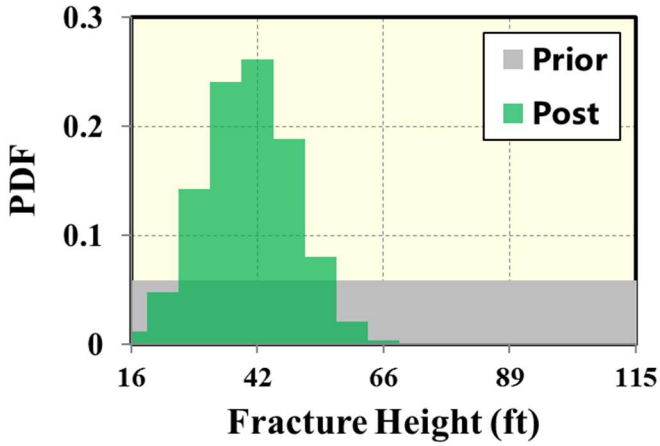


(e) Fracture water saturation

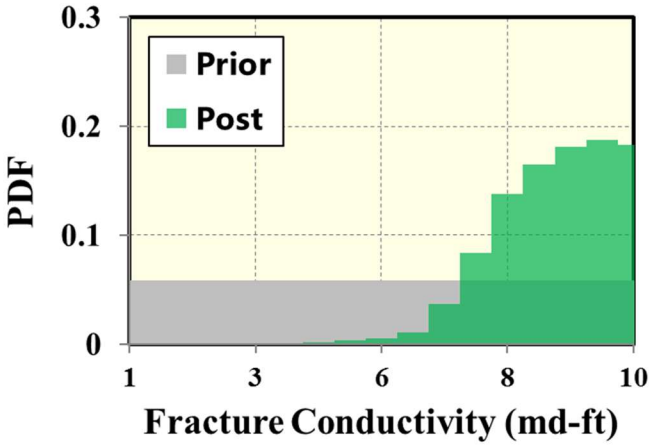
Figure 4.12: Comparison between prior and posterior probability density function of five parameters from 64 HM solutions: (a) Matrix permeability, (b) Fracture height, (c) Fracture conductivity, (d) Matrix water saturation and (e) Fracture water saturation.



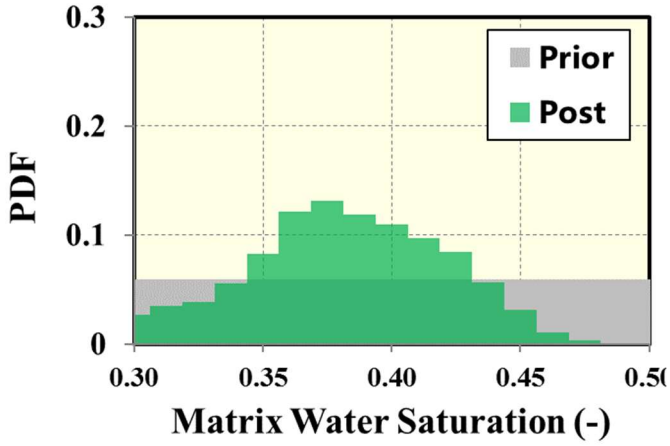
(a) Matrix permeability



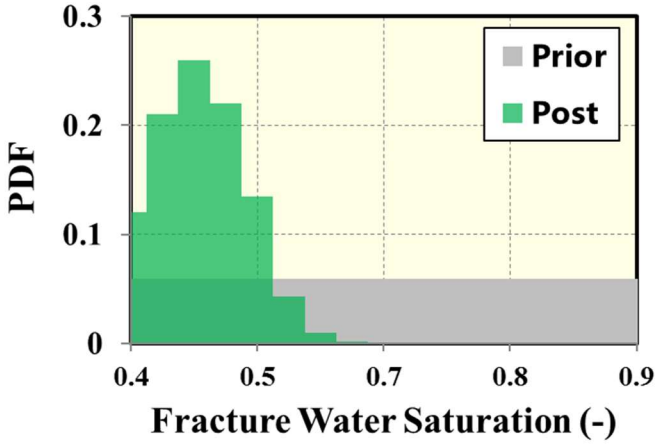
(b) Fracture height



(c) Fracture conductivity



(d) Matrix water saturation

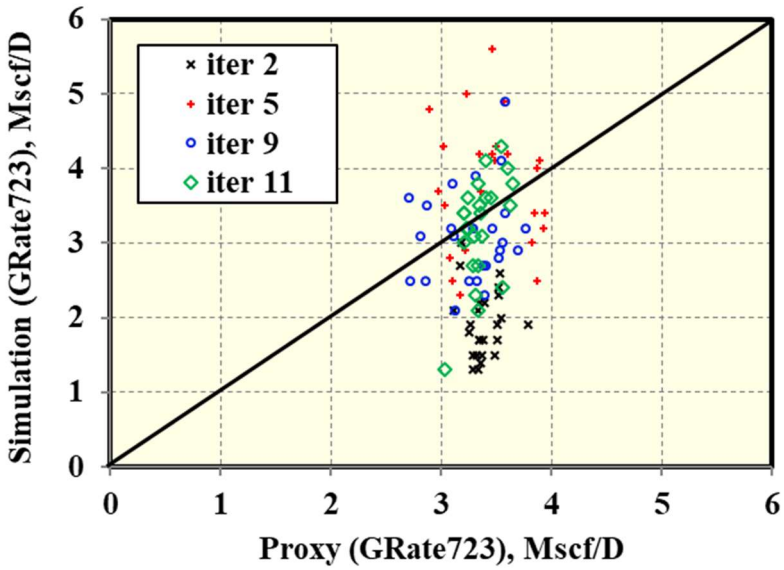


(e) Fracture water saturation

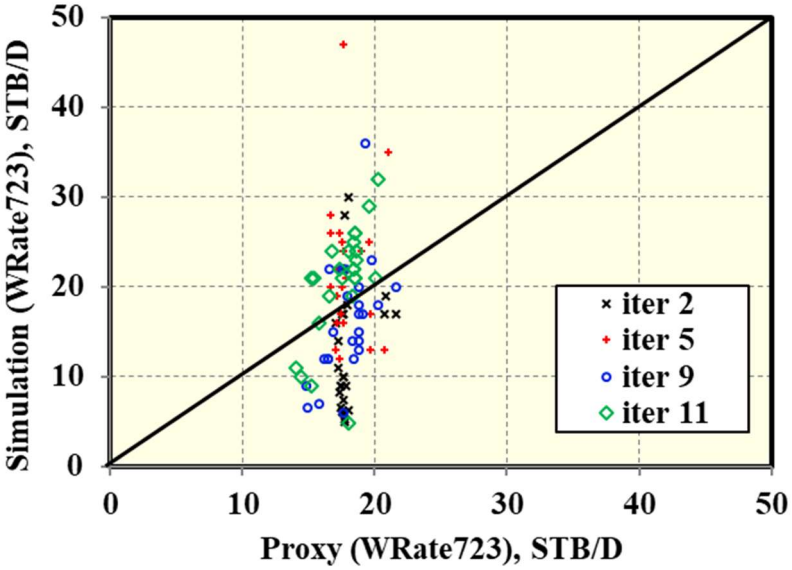
Figure 4.13: Comparison between prior and posterior probability density function of five parameters from 10000 HM solutions generated by proxy-based MCMC algorithm: (a) Matrix Permeability, (b) Fracture Height, (c) Fracture conductivity, (d) Matrix water saturation and (e) Fracture water saturation.

Figure 4.14a and 4.14b illustrate the improvement of proxy predictability in gas and water flow rate at day 723 (2 of 20 proxies). Despite the final proxy before stopping criteria triggered (iteration 11) is not perfect, we can still see the gradual improvement of proxy comparing to simulation results. The more iteration continues, the points are getting closer to unit slope line. Figure 4.14c also shows the objective function estimated from multiple-response-parameter proxies versus actual one from simulation runs. The improving trend is clearly seen when iteration progresses and there are more measured points for proxy modeling. One thing to be aware is that the objective of using proxy in HM is not to create the most accurate proxy to represent the highly non-linearity in actual simulation results. To achieve that, we have to actually run a lot more simulation cases as measured points. In statistics, there is a saying that most of proxies are wrong; some of them can only be used. As a result, the actual objective for using proxy is just to adopt proxy to guide where MH-MCMC algorithm should propose possible HM solutions in the iterative HM workflow while being aware of the inaccuracy of proxy.

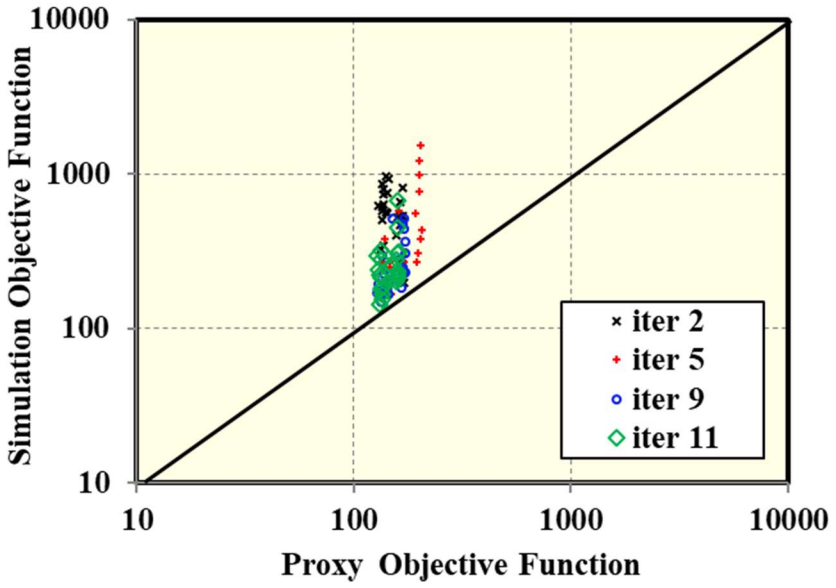
The workflow not only gradually improved proxy quality in the interested area in the iterative manner, but it also explored HM solutions at the same time (first, adding more points in the possible HM solutions domain, and second, adding some measured points by the farthest distance criteria from measured points to explore new possible HM solutions in other areas). Figure 4.15 shows the number of solutions versus iteration number. The number of solutions is getting higher at the later iterations. This is because the proxy accuracy at the later iteration is better due to having more measured points.



(a) Gas flow rate at day 723



(b) Water flow rate at day 723



(c) Objective function (global error)

Figure 4.14: Cross plot of simulation values and proxy-estimated values showing the improvement of proxy predictability as iteration progresses: (a) Gas flow rate at day 723, (b) Water flow rate at day 723, (c) Objective function (global error).

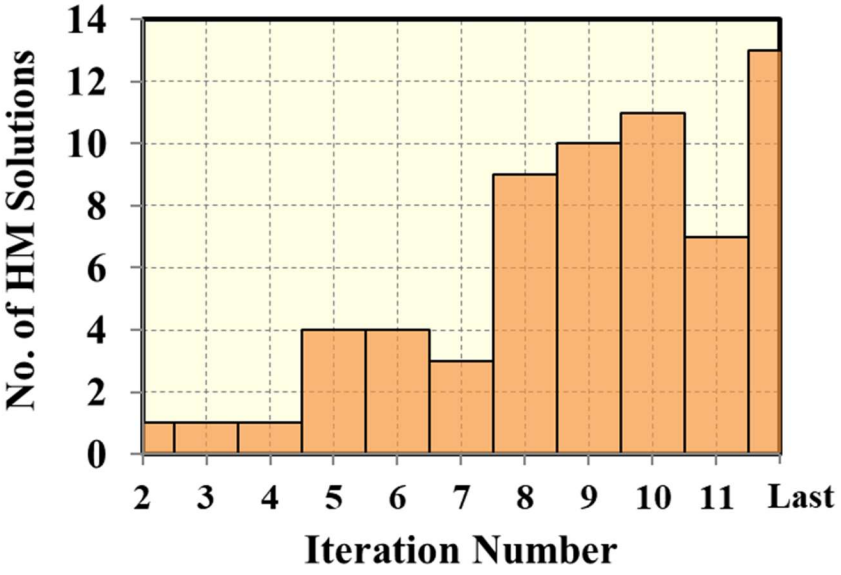
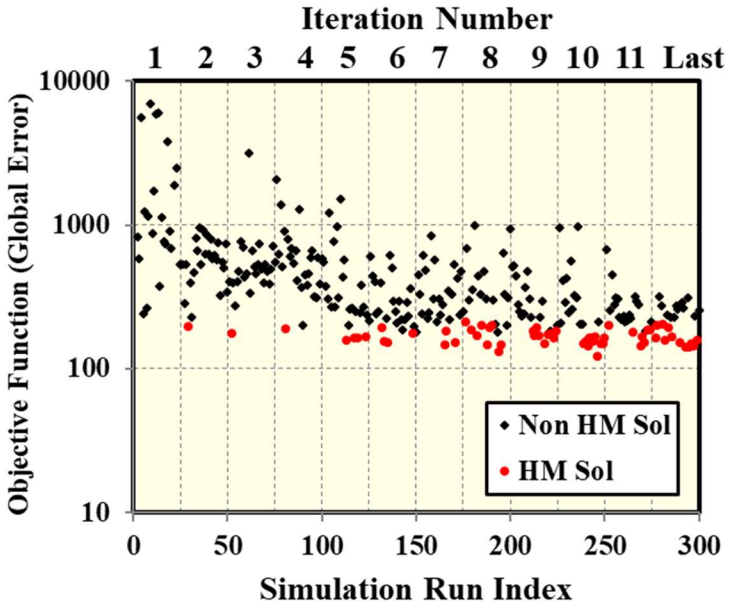


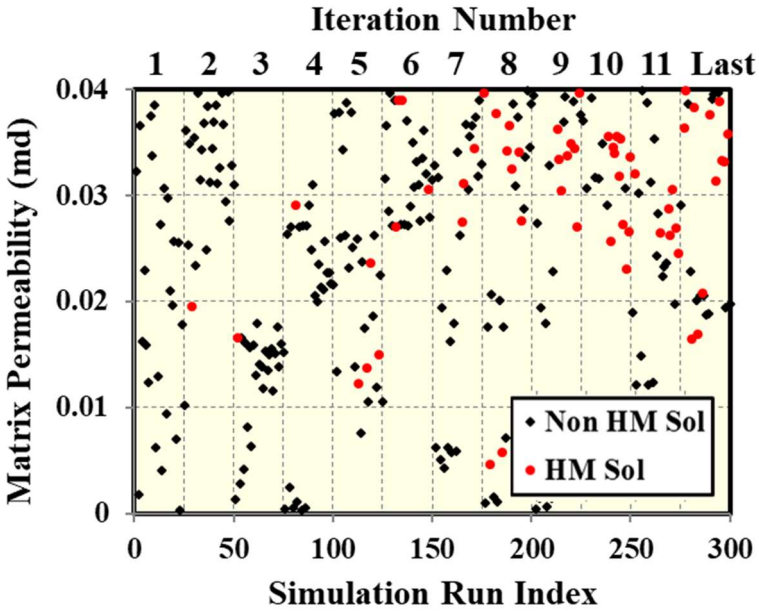
Figure 4.15: Number of HM solutions vs. iteration number.

Figure 4.16 shows objective function and each uncertain parameter of the proposed history realization cases vs. simulation run index. As shown, the objective function is improving when we have more simulation runs. This is to confirm the improvement of proxy modeling accuracy when we have more measured points. In the initial iteration where we used Latin Hypercube Design, we can see the simulation run cases scattered over the parameters domain. As iteration progresses and more simulation runs are available, the ranges of the proposed cases in each iteration are narrowed down to possible HM solutions domain for fracture height, fracture conductivity and fracture water saturation while the narrowing trend is less for matrix permeability and matrix water saturation. As mentioned, the proxy-based MCMC algorithm will sample and propose more points around HM solutions in order to improving proxy accuracy where we are interested in and also do not waste too much simulation runs in other area but still have them in order to find new possible HM solutions domain.

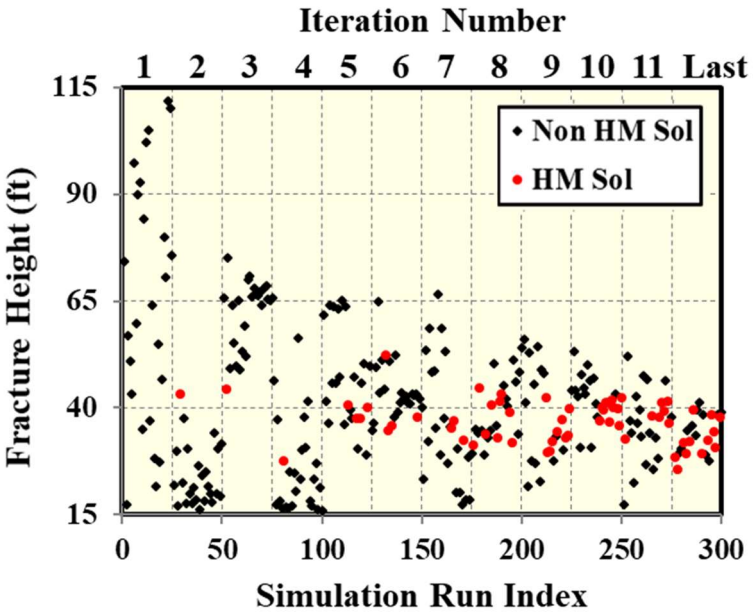
Figure 4.16 also shows that the proxy-based MCMC algorithm discovers more HM solutions (red dot) in the later iteration where proxies have more measured points. In addition, the algorithm still can find the new solutions in each new iteration and are not trapped in any local HM solutions as iteration progresses, especially for matrix permeability and matrix water saturation. In addition, we can see that the HM solutions in fracture height, fracture conductivity, fracture water saturation domain are quickly converged to specific ranges because these three parameters have more impact to the objective function. This confirms that proxy-based MCMC algorithm is efficient and practical to explore the HM solutions and to capture the uncertainty of HM realizations as posterior distribution.



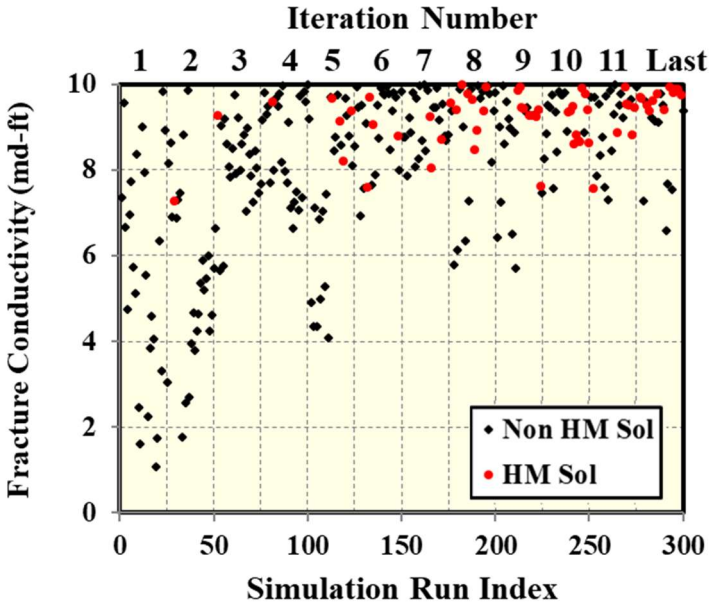
(a) Objective function (global error)



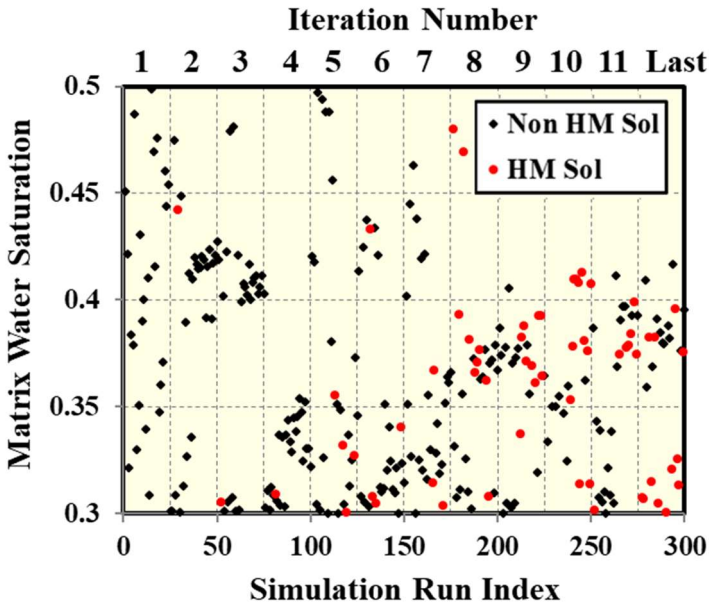
(b) Matrix permeability



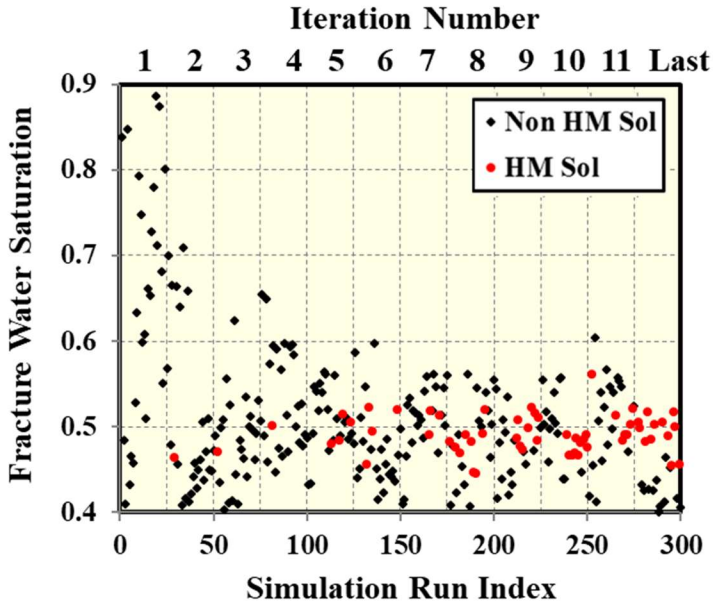
(c) Fracture height



(d) Fracture conductivity



(e) Matrix water saturation



(f) Fracture water saturation

Figure 4.16: Proposed points (HM solutions in red, non-HM solutions in black) vs. simulation run index and iteration number: (a) Objective function (global error), (b) Matrix permeability, (c) Fracture height, (d) Fracture conductivity, (e) Matrix water saturation and (f) Fracture water saturation.

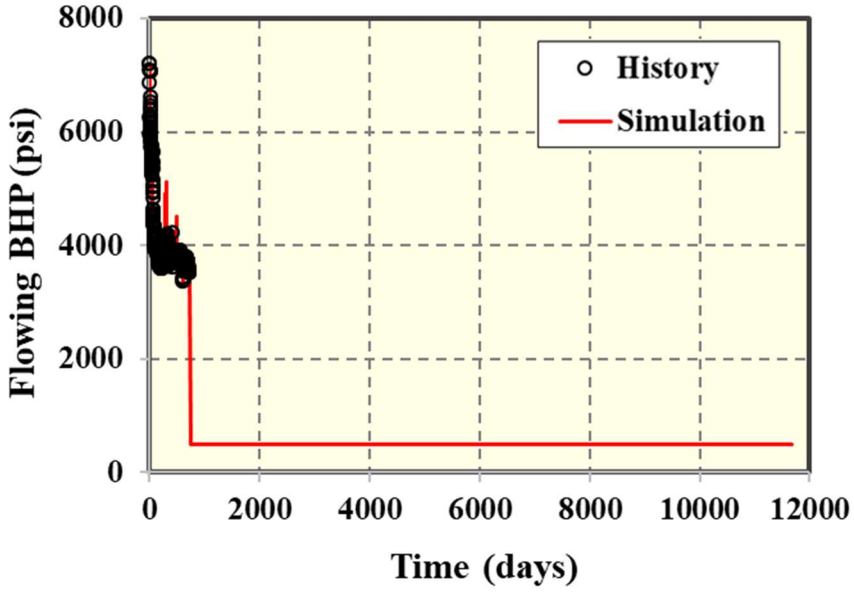
4.6 PROBABILISTIC FORECASTING

After obtaining HM solutions from the iterative HM workflow, we performed the production forecast. The code we developed will automatically generate the simulation files for production forecasting of all 64 HM solutions. We used BHP of 500 psi as production constraint. The production forecast period is 30 years. The results of gas rate, water rate, cumulative gas production and cumulative water production including the best match case and P10-50-90 cases are shown in Figure 4.17.

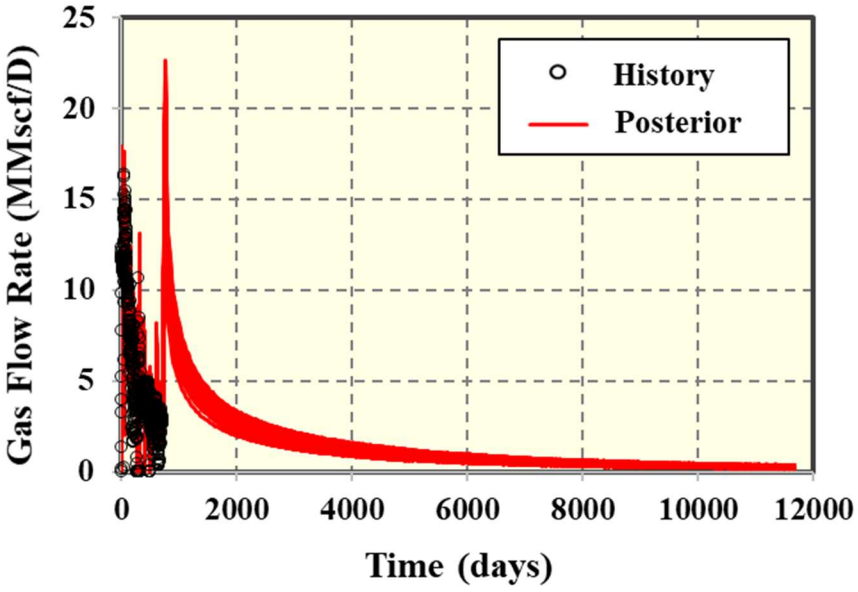
The Empirical Cumulative Distribution Function (ECDF) of gas Estimated Ultimate Recovery (EUR) from all 64 HM solutions is shown in Figure 4.18a. P10, P50 and P90 of gas EUR are 15.1, 18.8, and 20.3 Bscf, respectively while the gas EUR at end of history is 3.65 Bscf. The P90/P10 ratio of gas EUR is 1.34. The realization corresponding to P50 gas EUR has values for each parameter as follows: matrix permeability of 0.027 md, fracture height of 38 ft, fracture conductivity of 8.9 md-ft, matrix water saturation of 0.37 and fracture water saturation of 0.51. For the best match case, the gas EUR is 19.3.

To illustrate, we also generated 10000 HM solutions by the proxy-based MCMC algorithm and constructed gas EUR proxy to estimate gas EUR for all 10000 HM solutions. The comparison of ECDF of gas EUR between all 10000 HM solutions and 64 HM solutions is shown in Figure 4.18b. P10, P50 and P90 of gas EUR from 10000 HM solutions are 16.8, 19.4, and 20.7 Bscf, respectively. The P50 value from both 10000 HM solutions and 64 HM solutions is similar while the P10 and P90 of 10000 HM solutions are narrower. This difference is from the accuracy of gas EUR estimated from the gas EUR proxy. The gas EUR proxy predictability depends on the number of measured points, in this case 64 of actual HM solutions, and how they are distributed over uncertain parameter space. The computational time for probabilistic forecasting part is around 1 hour.

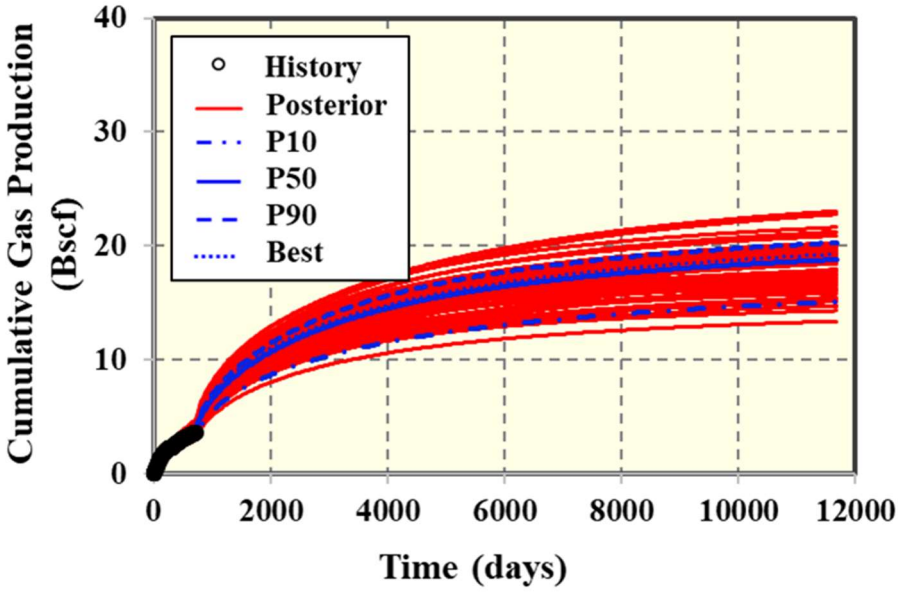
Finally, pressure distributions at different times are illustrated in Figure 4.19 in order to show reservoir depletion and the drainage area from one year of production until the end of production forecast of 30 years after historical period.



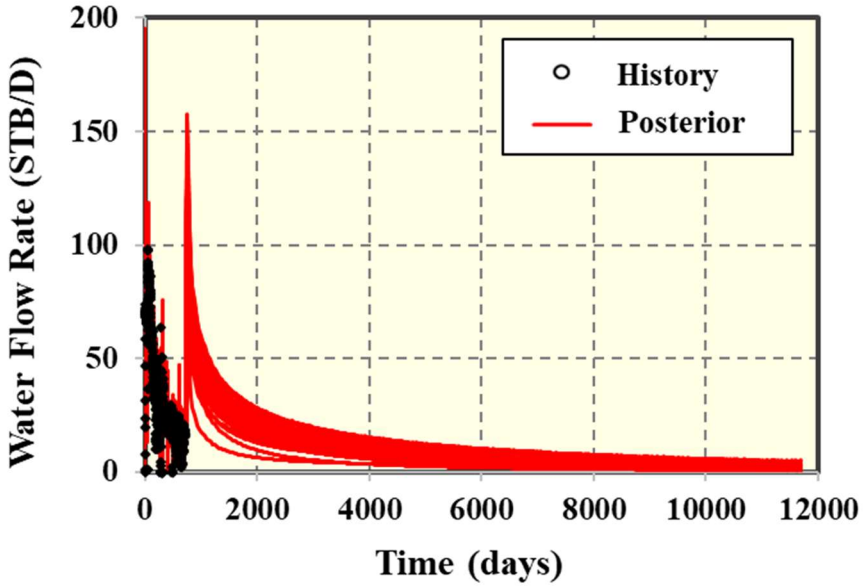
(a) Flowing BHP as constraint



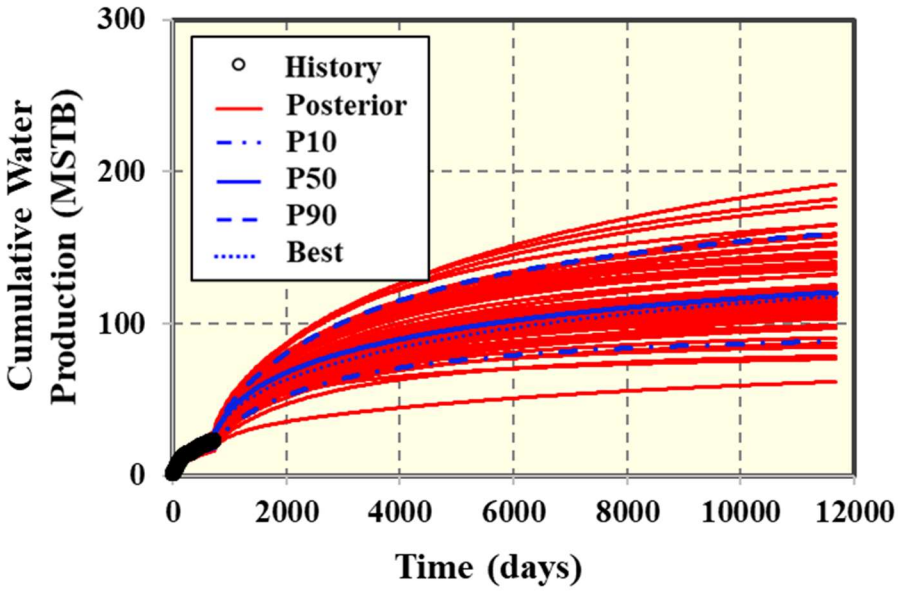
(b) Gas flow rate



(c) Cumulative gas production with P10, P50, P90 and the best match case

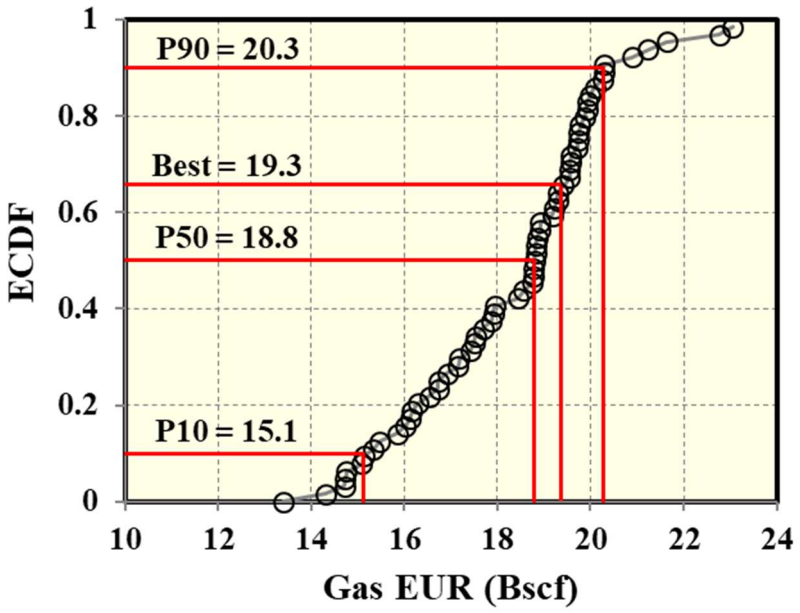


(d) Water flow rate

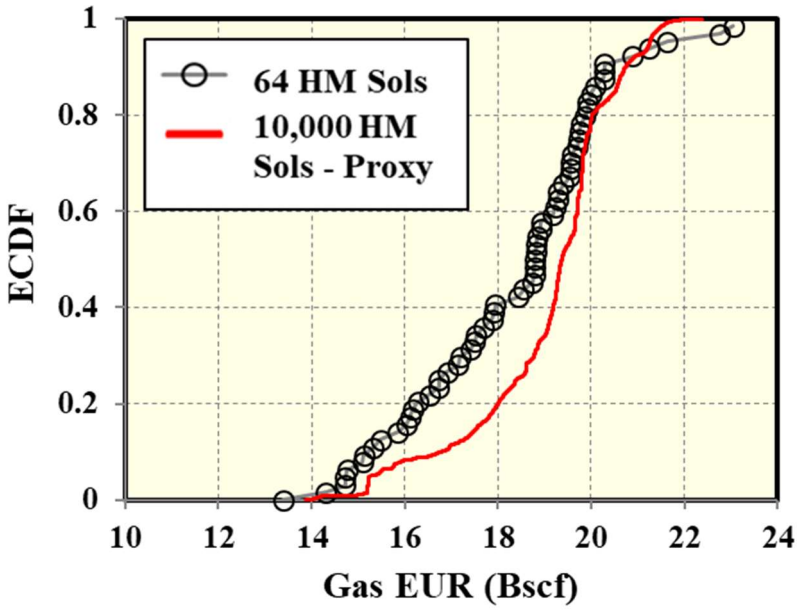


(e) Cumulative water production with P10, P50 and P90 and the best match case

Figure 4.17: Posterior production forecast of 64 HM solutions: (a) Flowing BHP as constraint (b) Gas flow rate, (c) Cumulative gas production with P10-50-90 and the best match case, (d) Water flow rate and (e) Cumulative water production with P10-50-90 and the best match case.



(a) 64 HM solutions



(b) Comparison between 64 HM solutions and 10000 HM solutions from proxy-based MCMC algorithm

Figure 4.18: ECDF of gas EUR: (a) 64 HM solutions and (b) Comparison between 64 HM solutions and 10000 HM solutions from proxy-based MCMC algorithm.

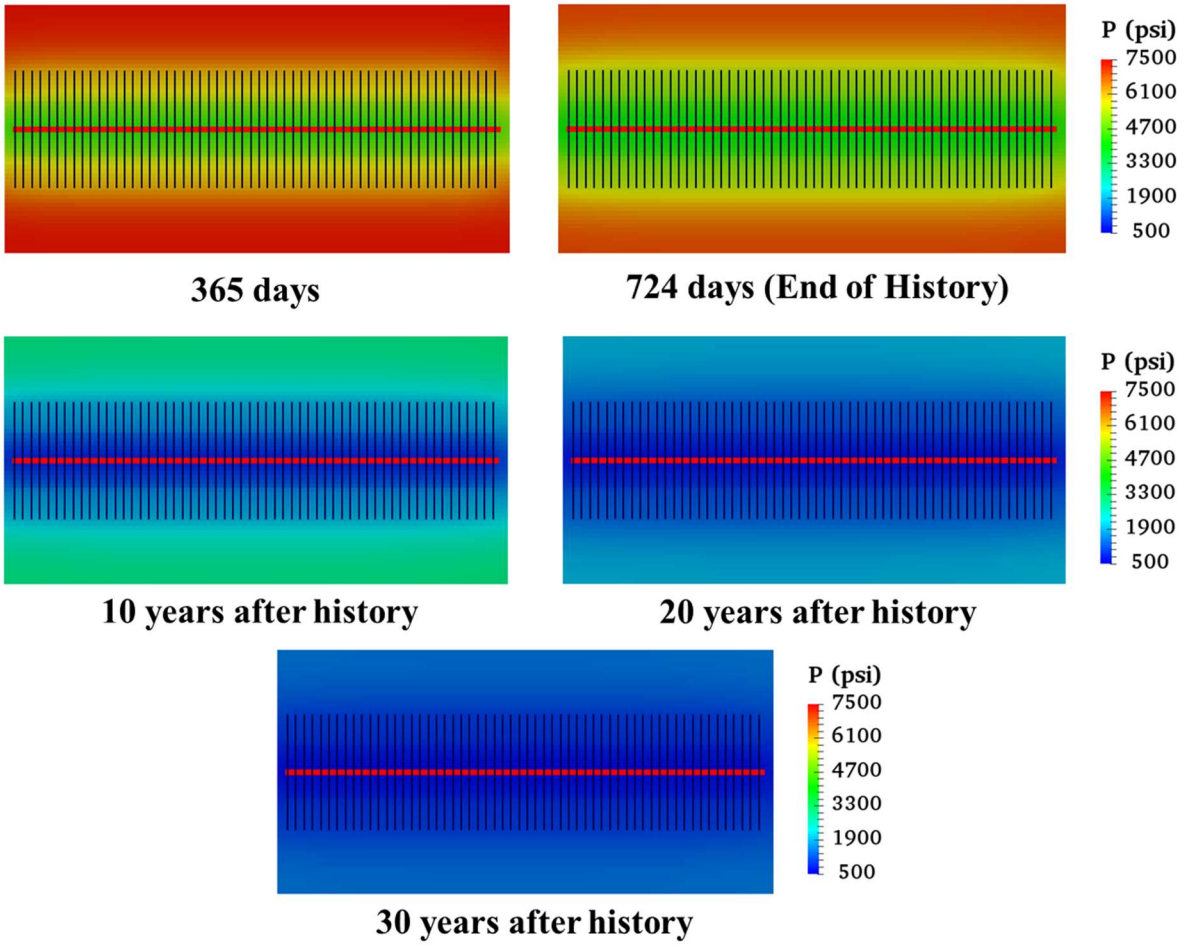


Figure 4.19: Pressure distributions of the best match case.

4.7 CONCLUSIONS

We developed a practical and efficient iterative AHM workflow through integrating the benefits of using multiple objective functions, multiple proxy-based MCMC algorithm, non-intrusive EDFM and commercial reservoir simulator. The workflow was applied to history match and perform probabilistic production forecast for a real case of shale gas well in China. In this chapter, we concluded as follows:

1. The AHM workflow efficiently performed history matching of gas rate and water rate for a shale gas well. The workflow found 64 HM solutions with 5 uncertain parameters from 11 iterations or 300 total simulation runs, which is around 21%. The computational time starting from parameter screening, history matching and probabilistic forecast is around 1 day.
2. The posterior distributions of 5 uncertain parameters were obtained from the AHM workflow. The reservoir and fracture properties uncertainties were effectively characterized and narrowed down.
3. The values with the highest frequency of each parameter are as follows: matrix permeability of 0.031 md, fracture height of 39 ft, fracture conductivity of 9.5 md-ft, matrix water saturation of 0.37 and fracture water saturation of 0.46.
4. The P50 gas EUR realization has properties as follows: matrix permeability of 0.027 md, fracture height of 38 ft, fracture conductivity of 8.9 md-ft, matrix water saturation of 0.37 and fracture water saturation of 0.51.
5. With the flowing BHP of 500 psi, the probabilistic production forecast of gas EUR in this shale gas well is 15.1, 18.8 and 20.3 Bscf for P10, P50 and P90, respectively.

Chapter 5: Assisted History Matching and Production Forecasting in Shale Gas Reservoirs Considering Enhanced Permeability Area

5.1 INTRODUCTION

Due to the fact that micro-fractures and fissures are normally created around main hydraulic fractures acting as simulated rock volume (SRV), modeling with only planar hydraulic fractures could be unrealistic and too simplified for this complex system flow behavior. Therefore, the objective of this chapter is to investigate how the Enhanced Permeability Area (EPA) around hydraulic fractures affects the history matching solutions comparing with the case with hydraulic fractures only. Also, we performed the comparison of different proxy models including K-nearest neighbors (KNN) and Neural Network (NN) about how each proxy performs in the multiple-proxy-based MCMC algorithm in terms of efficiency and running time.

All three main steps in the AHM workflow were performed including parameter identification and screening, history matching and production forecast. In this study, we used another shale gas well with 74-day historical production data. We had direct measurements of production data including gas and water flow rates while the bottomhole pressure (BHP) data were estimated from the wellhead pressure (WHP) with pressure-loss correlation. We structured this chapter by starting with the case with hydraulic fractures only comparing two different proxies including KNN and NN. Then, the case with EPA was performed with NN proxy and compared with the case with hydraulic fractures only.

5.2 RESERVOIR MODEL

The field case study is a shale gas well in China. The well is a horizontal well with a length of 4921 ft with 30 stages of hydraulic fractures and in each stage, 3 perforation clusters were shot. The total number of hydraulic fractures is 90.

We built a homogeneous reservoir model with a constant thickness. The model dimension is 5500 ft long, 2200 ft wide and 28 ft thickness. We also assumed hydraulic fractures penetrate the whole reservoir thickness, as model dimension in Z axis. The number of grid blocks is 275, 22 and 1 in x, y and z direction, respectively. Basic reservoir and fractures properties are summarized in Table 5.1. For shale gas reservoirs, we modeled gas desorption using Langmuir isotherm as shown in Table 5.2. Also, the effect of geomechanics during production depletion was also included. The pressure-dependent permeability for both matrix and fractures was modeled to capture the permeability reduction when reservoir pressure was depleted, by Equation 5.1. The permeability modulus of 0.055 and 0.056 were used for matrix and fractures, respectively. The normalized pressure-dependent permeability curves for both matrix and fracture are illustrated in Figure 5.1. The reservoir model including 90 hydraulic fractures in a horizontal well modeled by EDFM is demonstrated in Figure. 5.2.

$$k = k_i exp^{-\gamma \cdot (P_i - P)}, \tag{5.1}$$

where

k = pressure-dependent permeability (md)

k_i = permeability at initial reservoir pressure (md)

γ = permeability modulus (1/psi)

P_i = Initial reservoir pressure (psi)

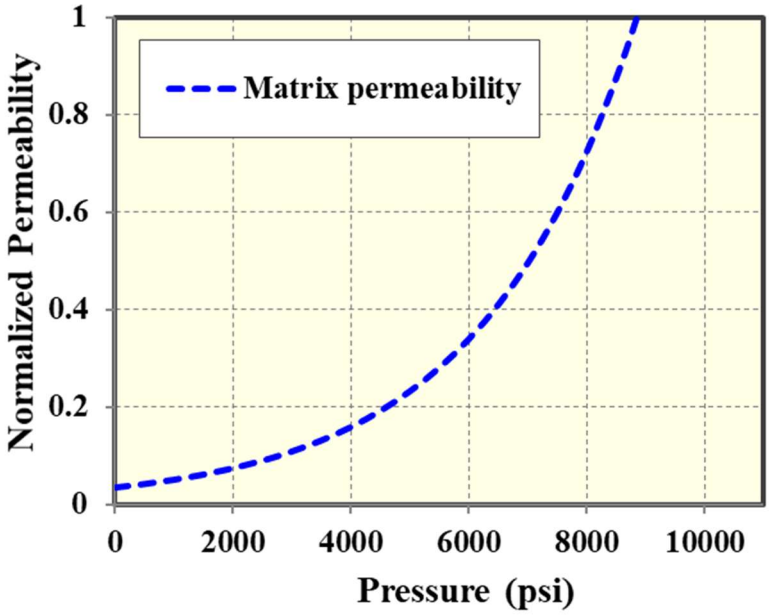
P = reservoir pressure (psi)

Reservoir description	Value	Unit
Model dimension (x × y × z)	5500 × 2200 × 28	ft
Number of grid blocks (x × y × z)	550 × 22 × 1	-
Grid block dimension (x × y × z)	10 × 100 × 28	ft
Initial reservoir pressure	8847	psi
Reservoir temperature	215	^{o}F
Residual water saturation	20%	-
Matrix water saturation	39%	-
Matrix porosity	5.6%	-
Total compressibility	3×10^{-6}	psi^{-1}
Reservoir depth	10499	ft
Well length	4921	ft
Number of stages	30	-
Clusters per stage	3	-
Cluster spacing	55	ft

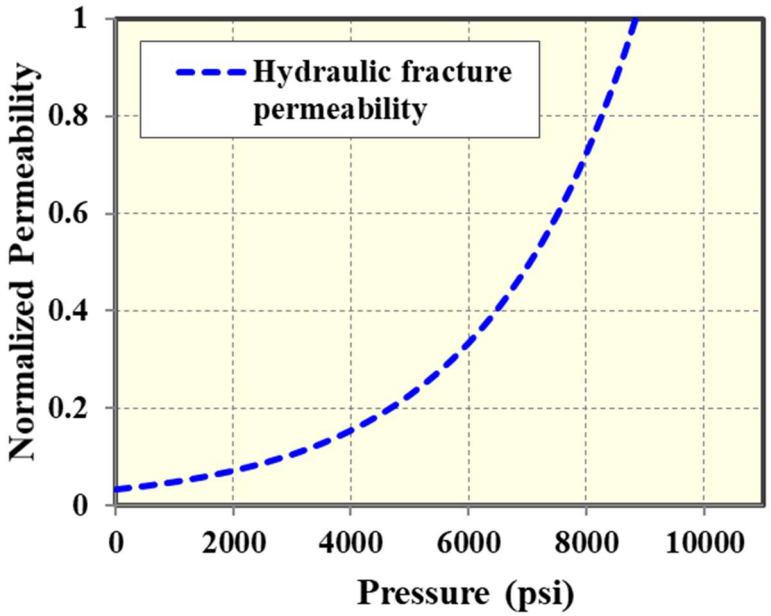
Table 5.1: Summary of basic reservoir and fracture parameters used in this study

Langmuir isotherm parameters	Value	Unit
Inverse-pressure parameter for the Langmuir isotherm model, CH_4	0.000985	1/psi
Maximum moles of adsorbed component per unit mass of rock, CH_4	0.071	gmole of component/lb of rock
Rock density	158.6	lb/ft^3

Table 5.2: Langmuir isotherm model parameters for the gas desorption effect



(a) Matrix permeability



(b) Fracture permeability

Figure 5.1: Pressure-dependent permeability for (a) Matrix permeability and (b) Fracture permeability to capture geomechanics effect of the permeability reduction when reservoir pressure was depleted.

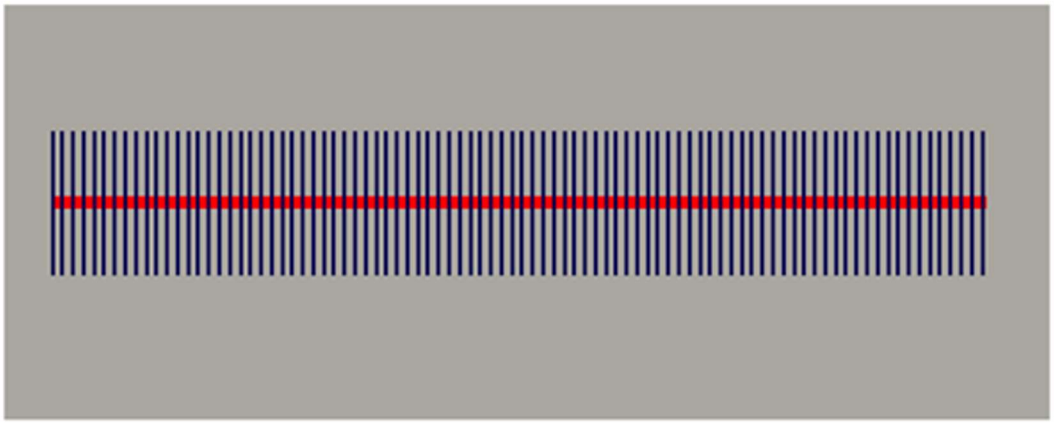


Figure 5.2: Reservoir model including 90 hydraulic fractures in a horizontal well.

5.3 Proxy Model Comparison between Neural Network (NN) and K-Nearest Neighbors (KNN)

To compare the performance between NN and KNN, we performed history matching twice using different proxy models. The steps are as AHM workflow including parameters identification and screening, history matching and production forecast as usual. The only difference is proxy type used in the proxy-based MCMC algorithm during the history matching.

5.3.1 Parameters Identification and Screening

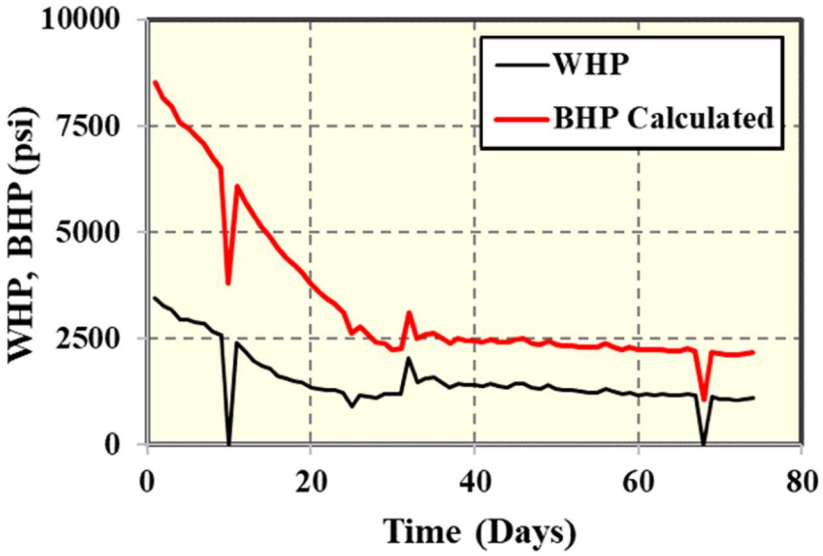
In this study, we identified 8 uncertain parameters and their ranges were summarized in Table 5.3 with uniform distribution based on prior knowledge about this well. The uncertain parameters can be grouped into two groups, fracture and matrix properties. We also included relative permeability curves parameters into uncertain parameters.

Code	Uncertain parameter	Unit	Distribution	Min value	Max value
A	Matrix permeability	mD	Uniform	0.00001	0.0001
B	Fracture half-length	ft	Uniform	200	1000
C	Fracture conductivity	md-ft	Uniform	1	100
D	Fracture water saturation	-	Uniform	0.6	0.9
E	Exponent of k_{ro}	-	Uniform	1	4
F	Endpoint of k_{rw}	-	Uniform	0.5	1
G	Exponent of k_{rw}	-	Uniform	1	4
H	Fracture width	ft	Uniform	0.01	5

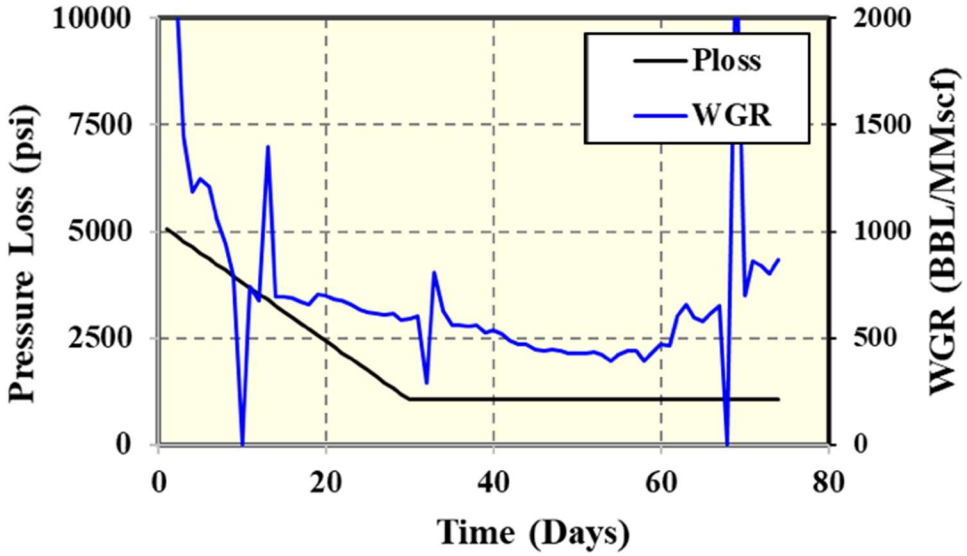
Table 5.3: Summary of eight uncertain parameters and their prior distributions

Then, we screened uncertain parameters for HM by performing partial two-level factorial design. For the screening purpose, we generated the 2^{8-2} or 64 cases to investigate each parameter effect and only partial interactions of parameters, not all interactions as full factorial design, which requires 2^8 or 256 cases. We constrained the production by gas flow rate and monitored BHP and water-gas ratio. The BHP is calculated from the available wellhead pressure and pressure loss estimation as shown in Figure 5.3a. The pressure loss is related directly to depth as calculated by Equation 5.2. However, to account for a high initial water gas ratio (WGR), we also added an extra pressure loss of 4000 psi at day 1, estimated from the pressure loss at initial water-gas ratio and at this well reservoir depth, and then linearly declining to 0 psi at day 30 as shown in Figure 5.3b.

$$BHP = WHP + 0.1\,psi/ft \times depth \tag{5.2}$$



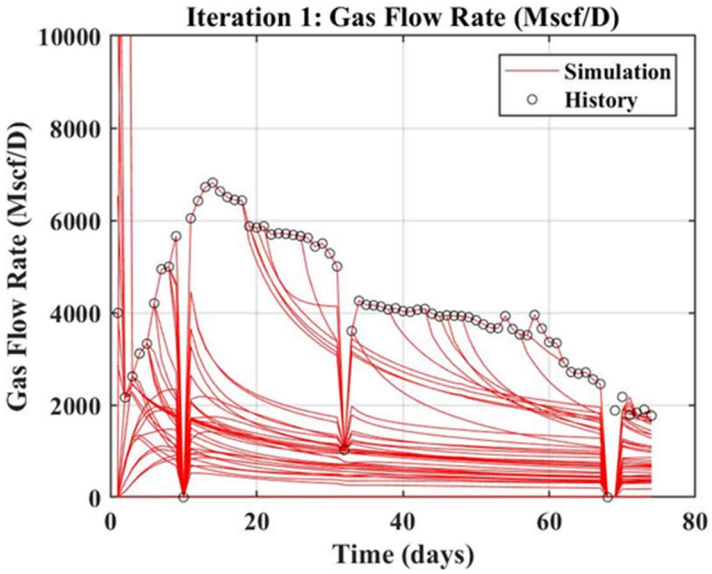
(a) WHP and calculated BHP



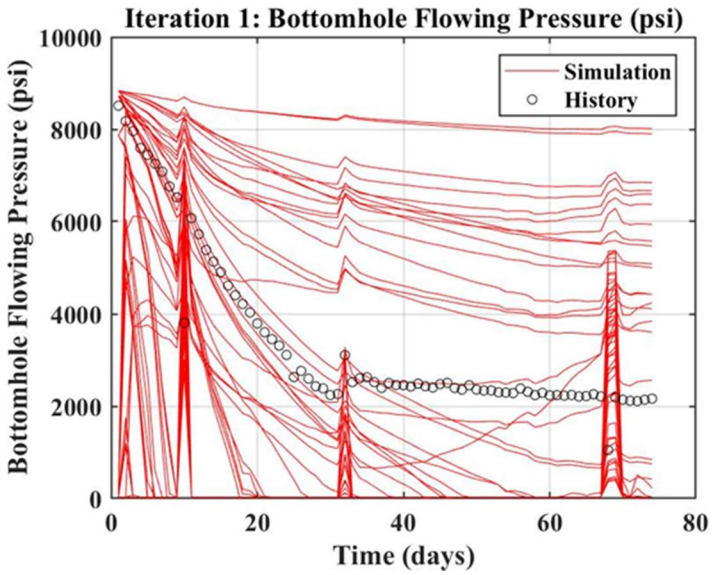
(b) Pressure loss and WGR

Figure 5.3: Calculate BHP from WHP by assuming pressure loss in the wellbore: (a) WHP and calculated BHP and (b) Pressure loss and WGR.

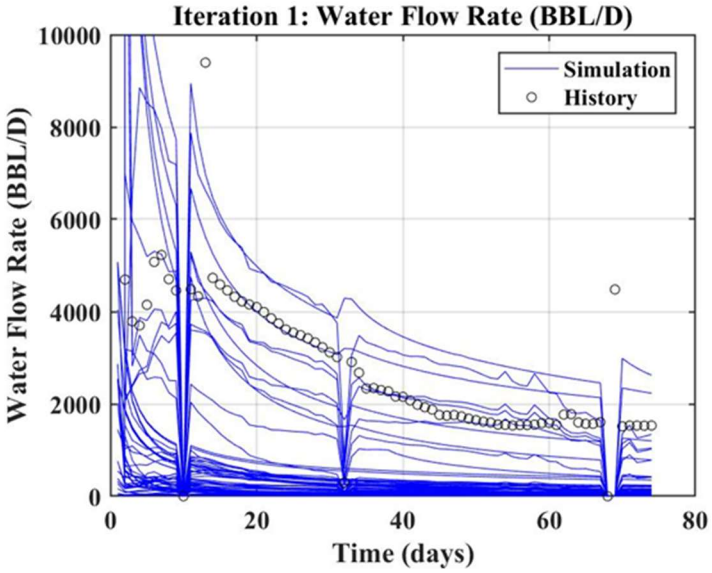
As in Chapter 4, to model hydraulic fractures, we used Embedded Discrete Fracture Model (EDFM) preprocessing code coupling with commercial simulator to create simulation files automatically. The simulation results of 64 cases are shown in Figure 5.4. One thing to check is whether the simulated data covered the historical data or not. If they did not cover the historical data, the initial ranges needed to be revised. Nevertheless, in this case, the results confirm the specified ranges of 8 uncertain parameters are acceptable and do not need the revision of uncertain-parameter range at this step.



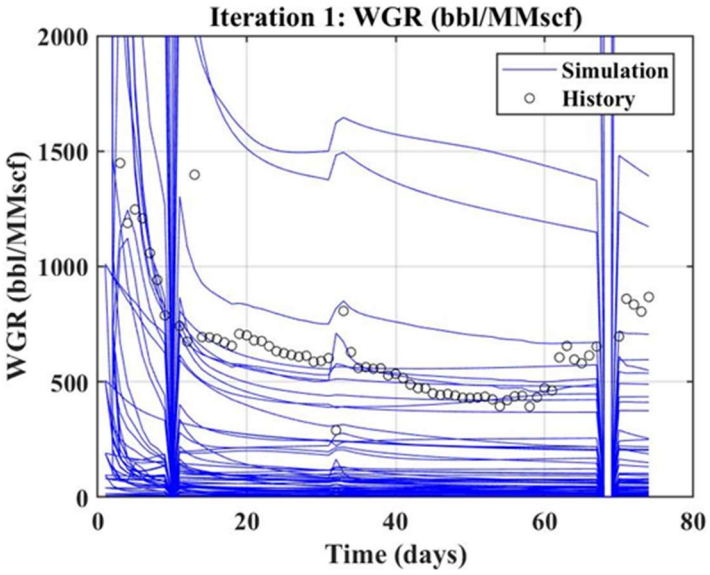
(a) Gas flow rate as constraint



(b) Flowing BHP



(c) Water flow rate



(d) WGR

Figure 5.4: Comparison between simulation results and historical data of 64 cases from partial two-factorial design: (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) Water-gas ratio (WGR).

Then, we performed t-value test or f-value test to see what uncertain parameters have more impact on BHP and WGR. In this study, we used f-value test and the plots of F-values of BHP error response and WGR error response of the individual parameters are shown in Figure 5.5 and 5.6, respectively. The formula to calculate response error is Equation 3.18.

For the BHP error response, B, C, D and H that are fracture half-length, fracture conductivity, fracture water saturation and fracture width, have more effect to the response. For the WGR error response, C, D, E, and H, which are fracture conductivity, fracture water saturation, exponent of k_{rg} and fracture width have more impact than the others. By considering both responses, we could screen from 8 parameters to 5 parameters by not

including matrix permeability, exponent of k_{rw} and endpoint of k_{rw}. However, for a later demonstration, we will include all 8 uncertain parameters to the history matching part.

Because we used multiple proxies instead of a single proxy for each error response to lessen the non-linearity of the problem; therefore, we had to identify the points to create multiple proxy models. We selected ten proxy points, five proxy points for BHP at day 7, 23, 36, 55 and 72, and another five proxy points for WGR 7, 17, 28, 40 and 59 as shown in Figure 5.7. Adding more proxy points should yield more valuable information such as different production trend. In this case, the selected 10 proxy points can capture the major trend of production data.

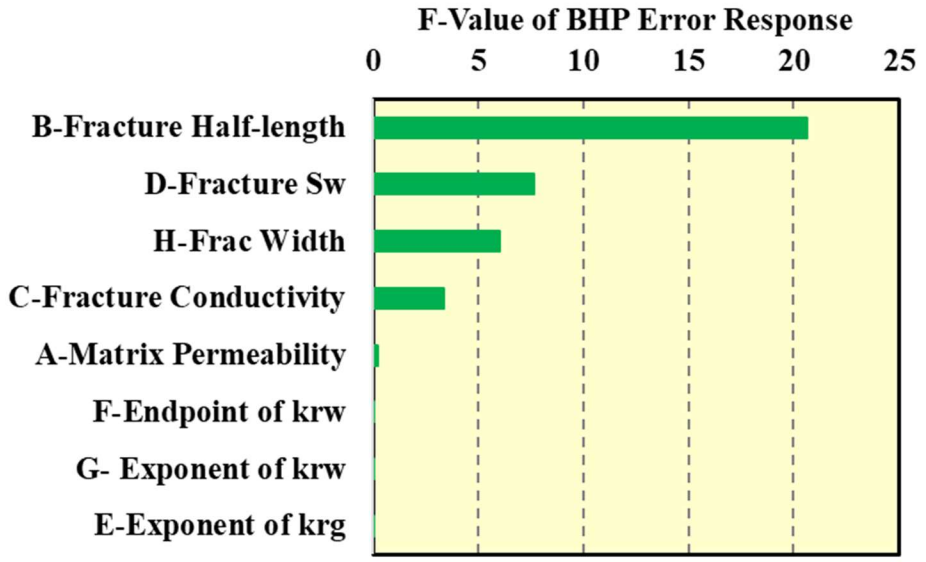


Figure 5.5: F-value ranking of partial two-level factorial design for the BHP error response.

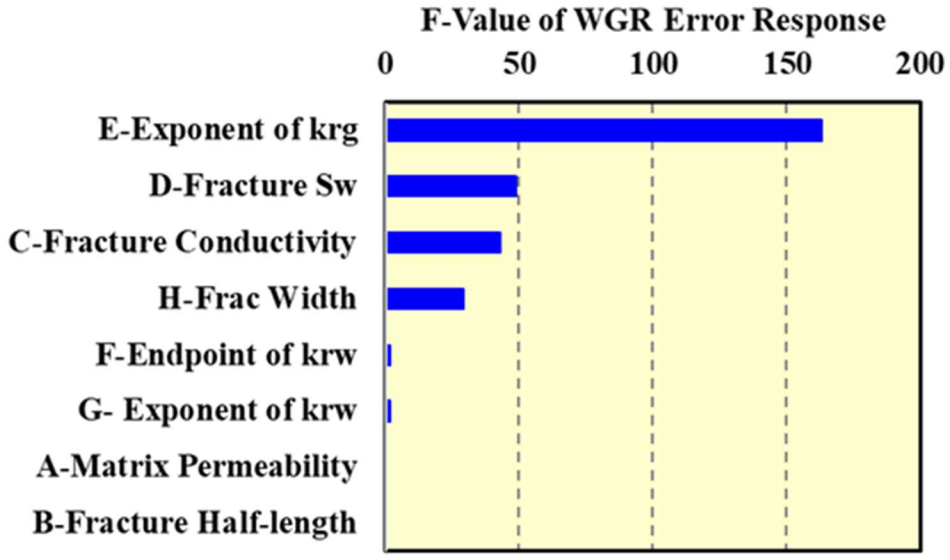
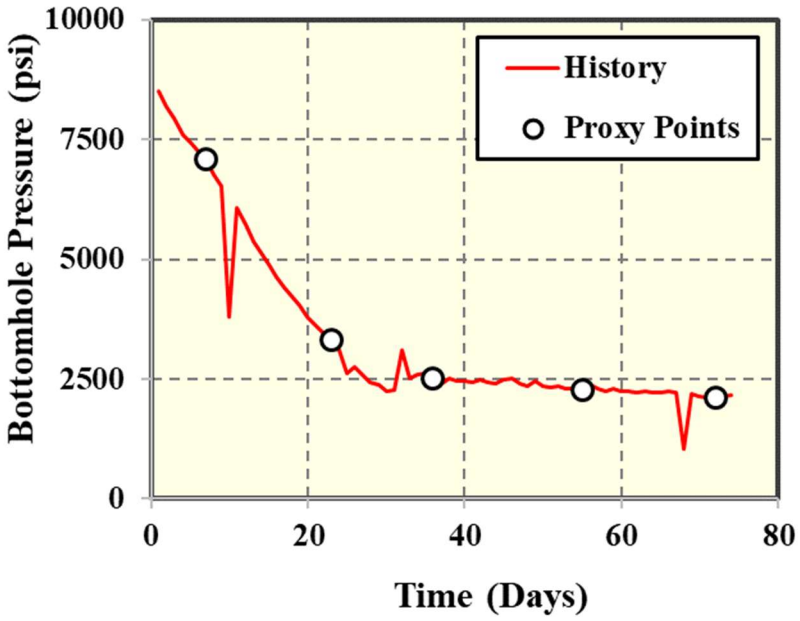
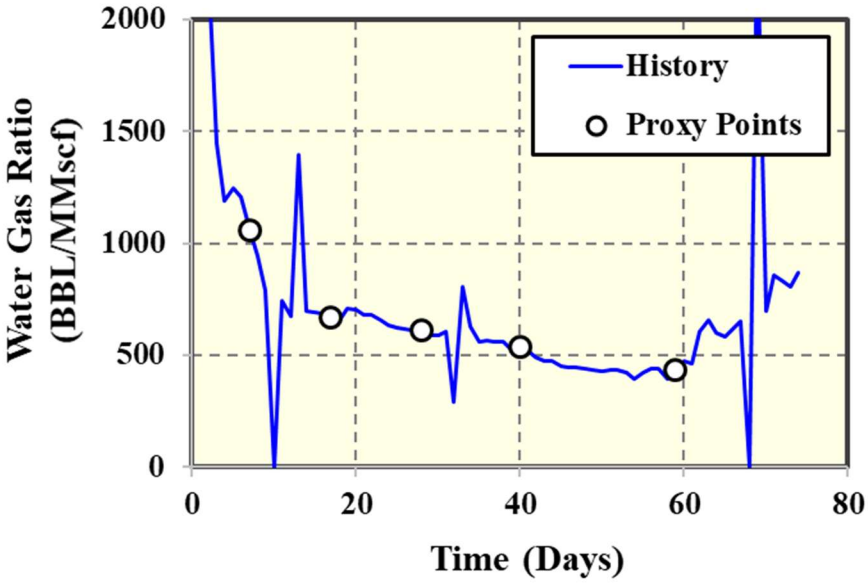


Figure 5.6: F-value ranking of partial two-level factorial design for the WGR error response.



(a) 5 proxy points of BHP

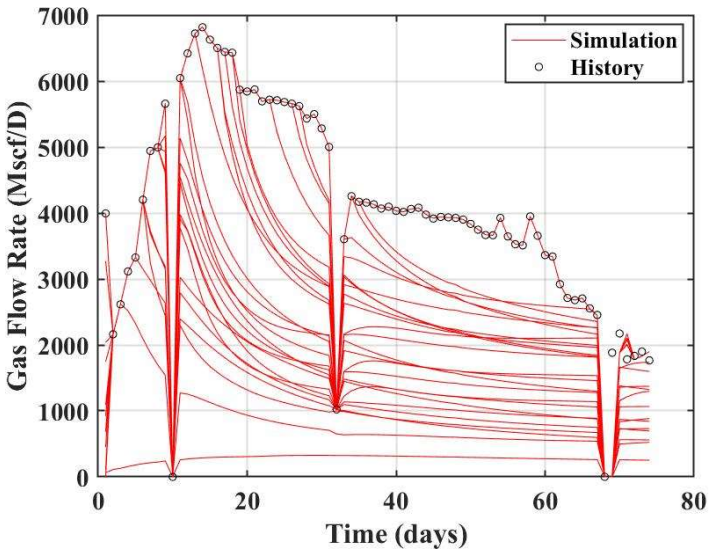


(b) 5 proxies of WGR

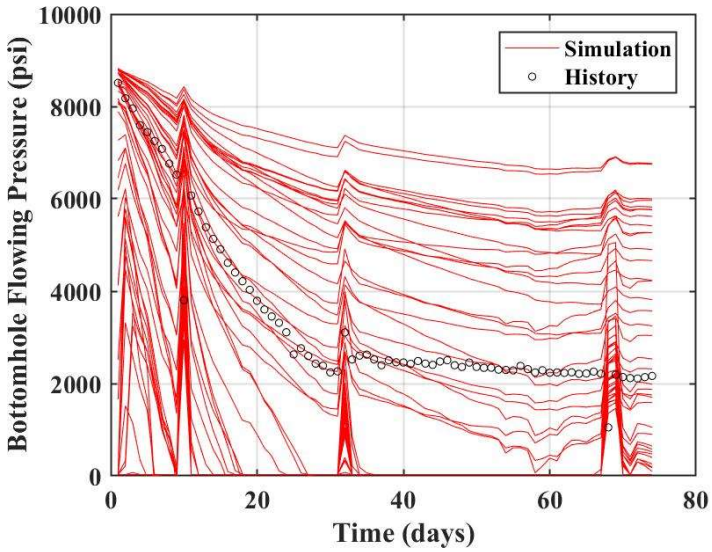
Figure 5.7: 10 multiple proxies including (a) 5 proxies of BHP and (b) 5 proxies of WGR.

5.3.2 History Matching

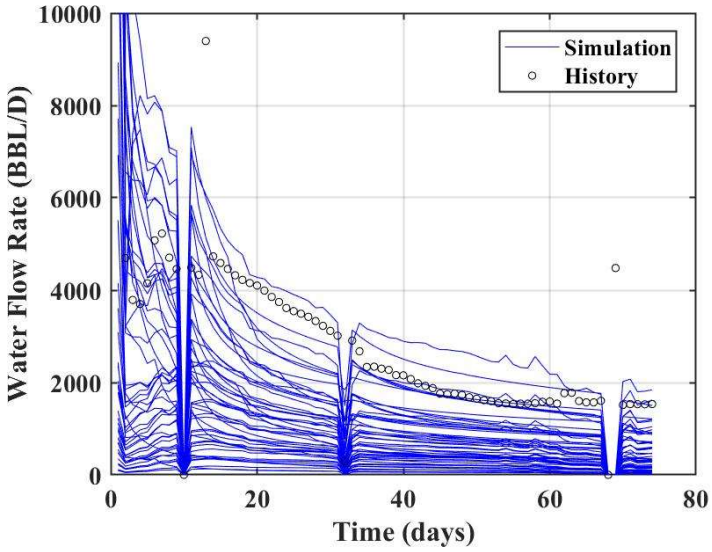
After we obtained uncertain parameters to be used for history matching, we generated initial iteration cases by using Latin Hypercube Design to sample the initial 50 cases from 8 uncertain parameters summarized as shown in Table 5.3, scattering over the uncertain-parameter domain. Figure 5.8 shows the results of initial 50 cases from Latin hypercube design. The simulation results also cover the historical production data as two factorial design but they distributed more evenly this time. With distributing more uniformly, it will be beneficial to construct multiple proxy models in the next step either NN or KNN. These multiple proxies will be later used in the proxy-based MCMC algorithm to propose new possible realizations as HM solutions in the next iteration.



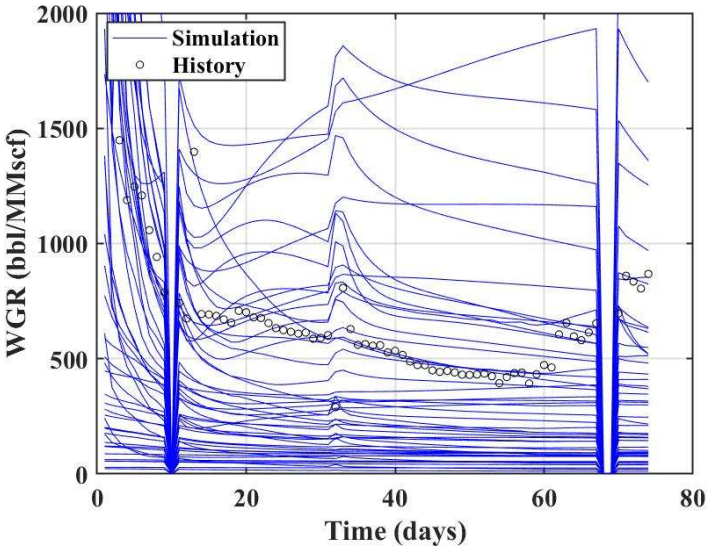
(a) Gas flow rate as constraint



(b) Flowing BHP



(c) Water flow rate



(d) WGR

Figure 5.8: Comparison between simulation results and historical data of 50 cases from Latin hypercube design: (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) WGR.

Then, the iterative HM workflow continued automatically by the developed code. We used one chain of MCMC and started the chain at the lowest global error points at each iteration. The proxy-based MCMC algorithm with filtering scheme proposed the next 25 cases for the next iteration to be validated with reservoir simulator. Then, these actual reservoir simulation results in each iteration were included into the training data set to further train proxies and improve the quality of proxies.

After that, the AHM workflow progressed until the stopping criteria was triggered, either relative discrepancies between the current and previous proxy with a threshold of 5% or the maximum iteration number (12 iterations or 325 simulation runs). For this study, the workflow stopped at the maximum iteration of 12 for both NN and KNN.

Next, we screened all cases from all iterations with the history matching solution criteria of 20% for both BHP and WGR error. The algorithms found 111 HM solutions (34%) and 20 HM solutions (6%) from a total 325 simulation runs for NN and KNN, respectively, as shown in Figure 5.9. We can observe that NN performs much better in terms of proposing higher number of history matching solutions. One thing to point out is that more solutions are found in the later iteration because the proxy model was improved by having more cases to train the proxy models in the interested domain.

With the low number of history matching solutions for k-nearest neighbors, we continued running until the number of history matching solutions is approximately the same as ones from NN for the comparison reason. In this case, we ran until 30 iterations and we obtained 105 history matching solutions from 775 simulation runs (14%) as shown in Figure 5.10. the simulation results of history matching solutions for NN and KNN are plotted in Figure 5.11-5.12.

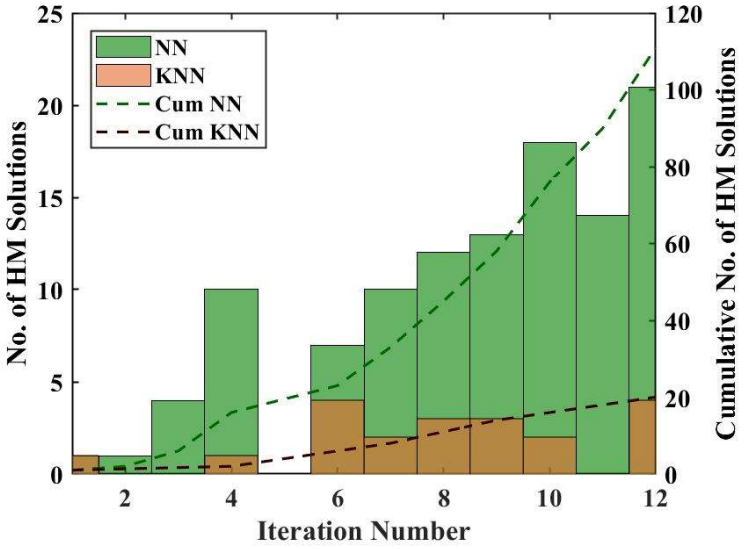


Figure 5.9: Comparison of number of history matching solutions vs. iteration number and between NN and KNN until iteration 12.

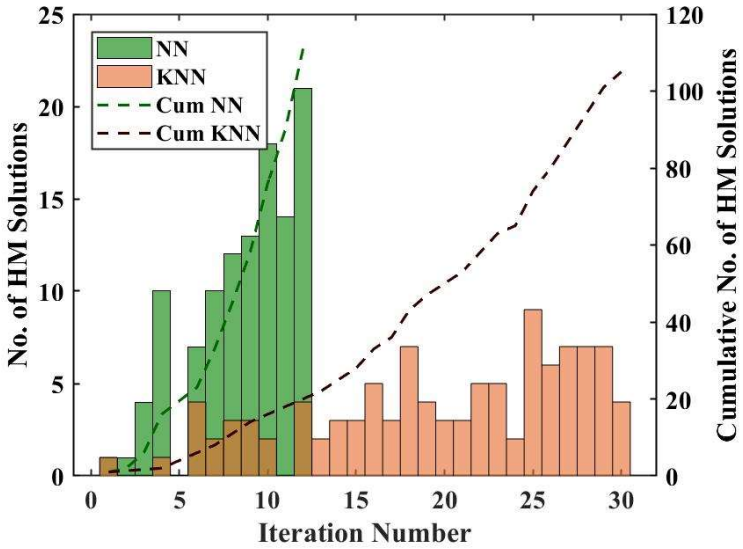
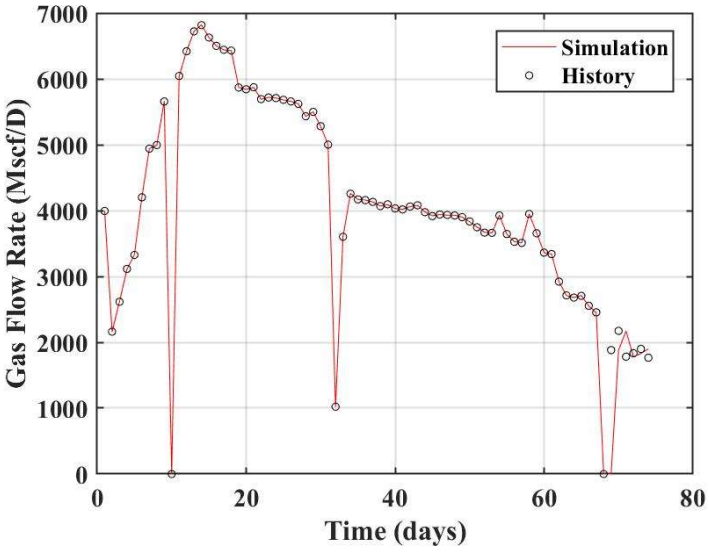
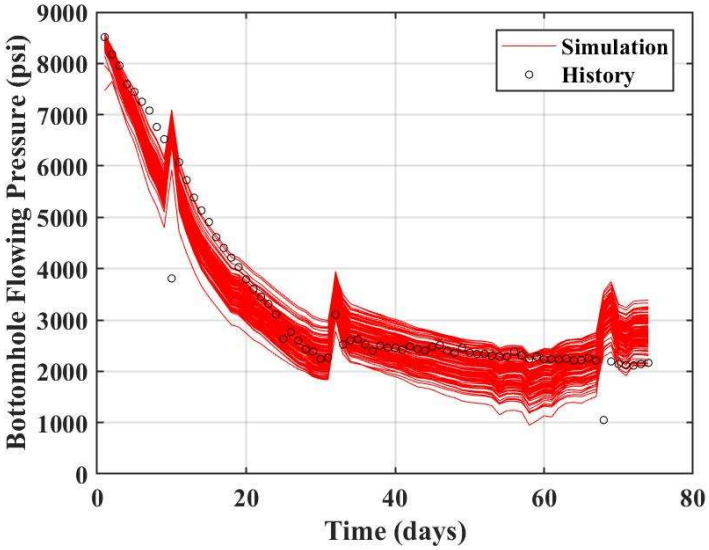


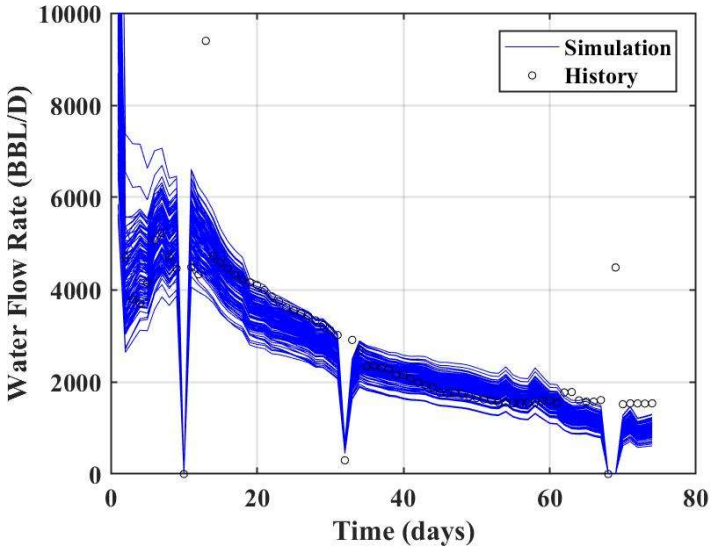
Figure 5.10: Comparison of number of history matching solutions vs. iteration number and between NN (until iteration 12) and KNN (continue running until obtain around 100 history matching solutions).



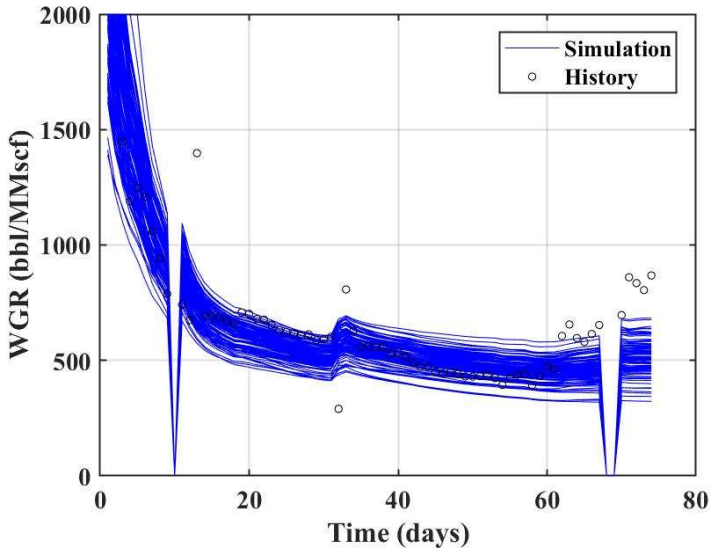
(a) Gas flow rate as constraint



(b) Flowing BHP

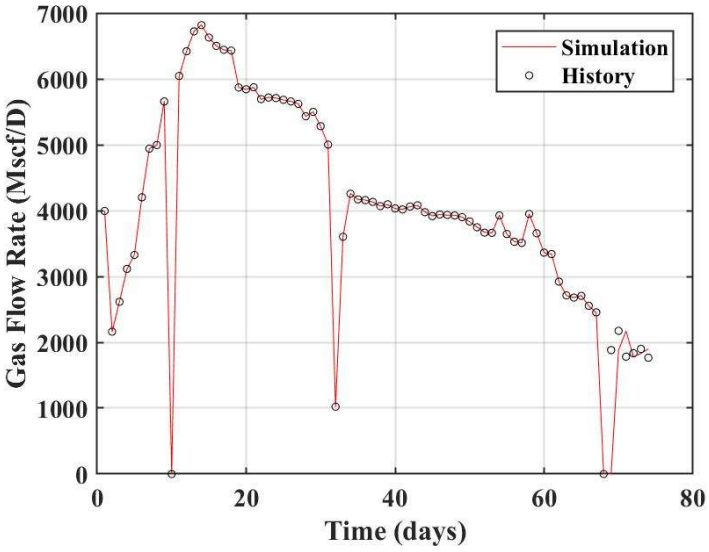


(c) Water flow rate

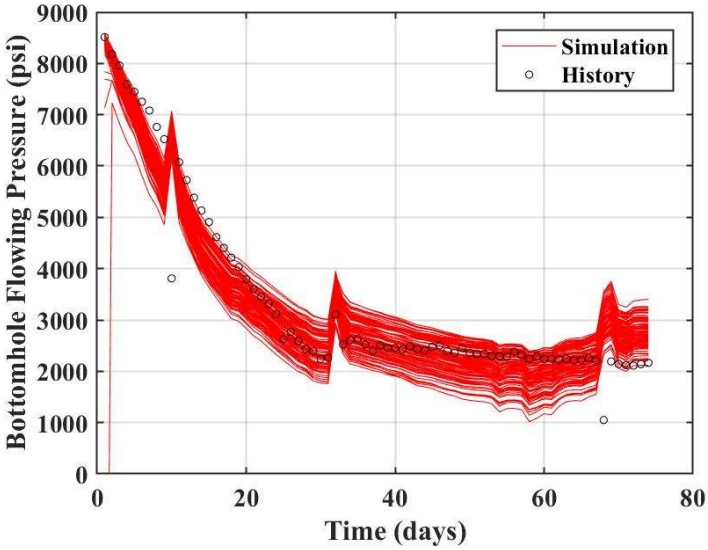


(d) WGR

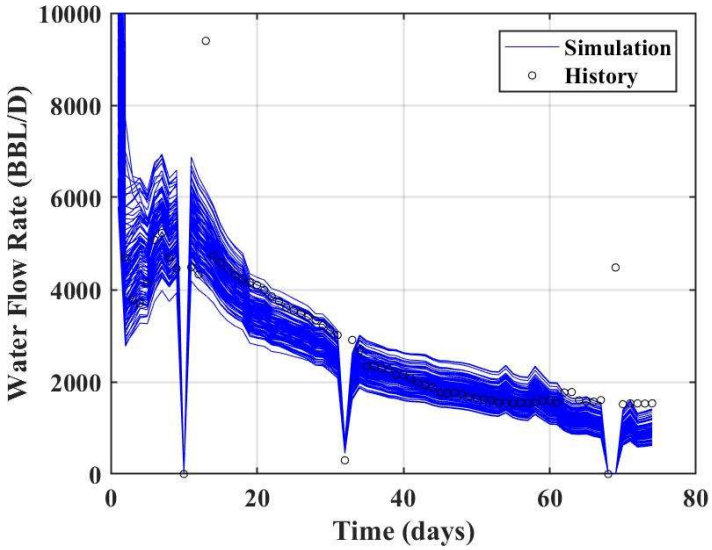
Figure 5.11: Comparison between simulation results and historical data of 111 history matching solution of using NN as proxy model: (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) WGR.



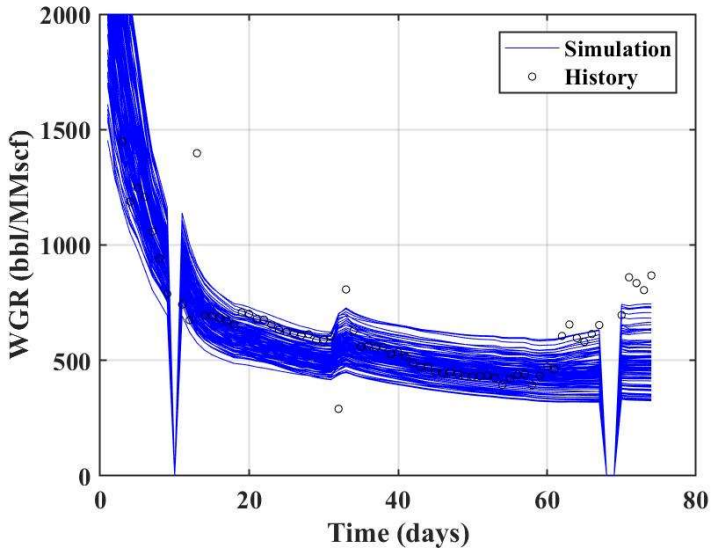
(a) Gas flow rate as constraint



(b) Flowing BHP



(c) Water flow rate

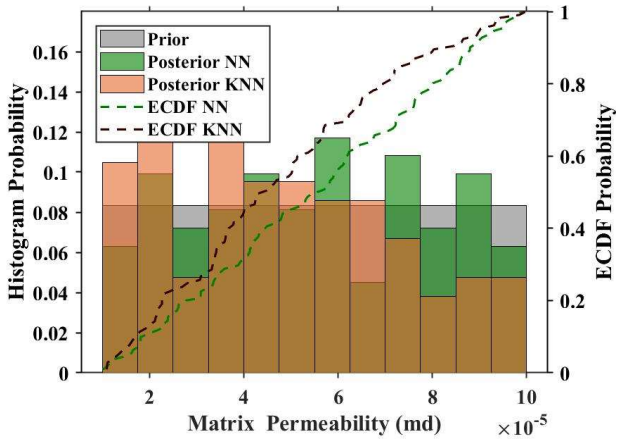


(d) WGR

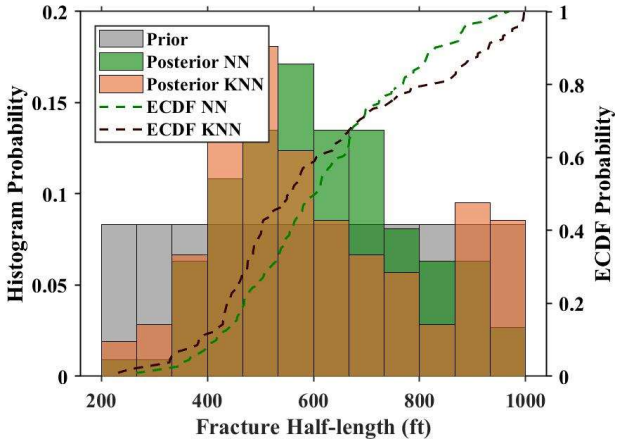
Figure 5.12: Comparison between simulation results and historical data of 105 history matching solution of using KNN as proxy model: (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) WGR.

Next, we could obtain the posterior distribution of history matching solutions from the AHM workflow for both NN and KNN as shown in Figure 5.13. Overall, both NN and KNN provide similar posterior distributions of each uncertain parameter. Posterior distribution of fracture half-length, fracture conductivity, fracture water saturation and fracture width have changed significantly from prior distribution; for example, higher side of fracture conductivity above 50 md-ft, fracture half-length around 400-700 ft, fracture water saturation around 0.75-0.85 and higher side of fracture width above 2 ft. On the other hand, the posterior distribution of uncertain parameter such as matrix permeability changed less significantly from prior distribution and any values of matrix permeability can provide history-matching solutions. These results confirm the screening step of identifying insignificant parameters such as matrix permeability. In sum, the results of in-lined posterior distribution between NN and KNN, and changes in posterior distribution guarantee the accuracy and efficiency of the proxy-based MCMC algorithm that we used as history matching algorithm.

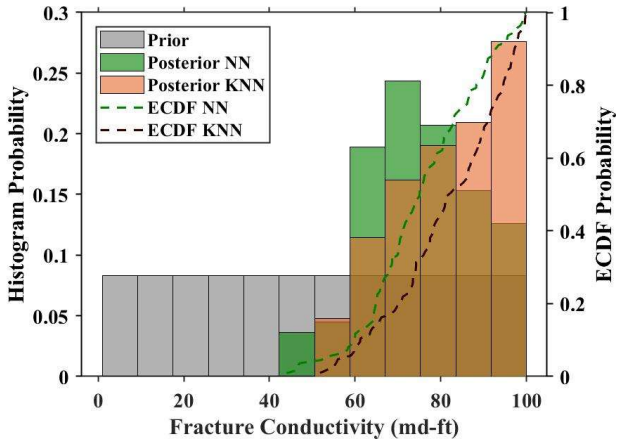
To obtain smoother posterior distributions, we generated 1000 solutions from proxy models for both NN and KNN as shown in Figure 5.14. However, one thing to be highlighted is that these 1000 solutions are not validated with the actual reservoir simulator, this means some cases will not be history-matching solutions when they validate with the simulator. This can be illustrated by Figure 5.10 that KNN has a low accuracy rate about 5 solutions out of 25 proposed cases at iteration 30 while NN has a high accuracy rate of 21 solutions out of 25 proposed cases at iteration 12. With this reason, as shown in Figure 5.14h for fracture width, we can observe the posterior distribution difference of proxy solutions between NN and KNN. However, for other uncertain parameters, there is no significant difference between NN and KNN for proxy solutions.



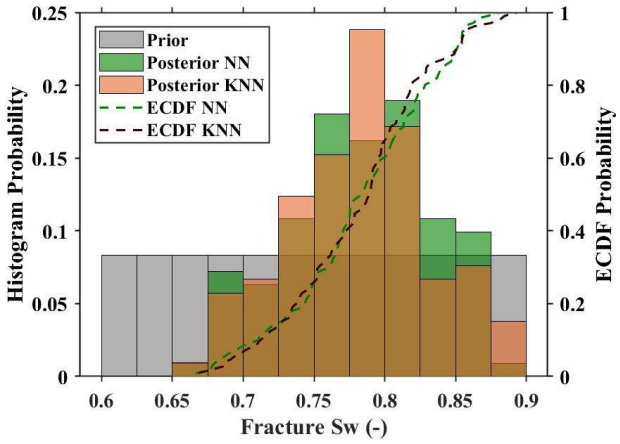
(a) Matrix permeability



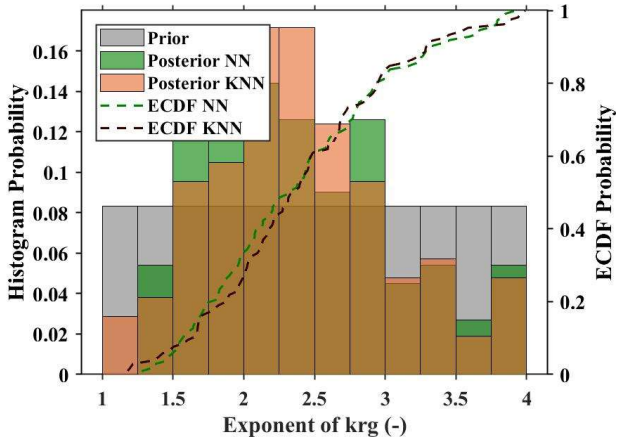
(b) Fracture half-length



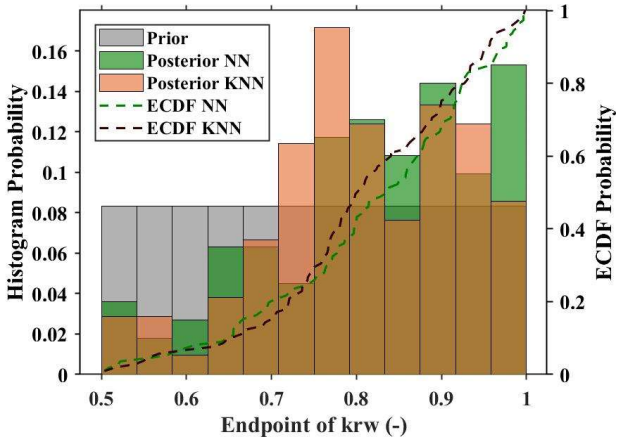
(c) Fracture conductivity



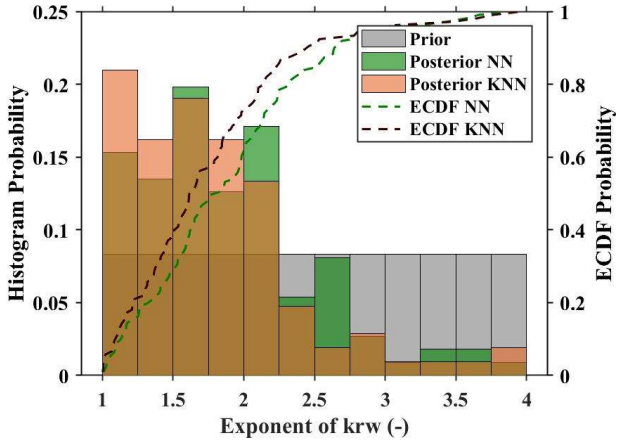
(d) Fracture water saturation



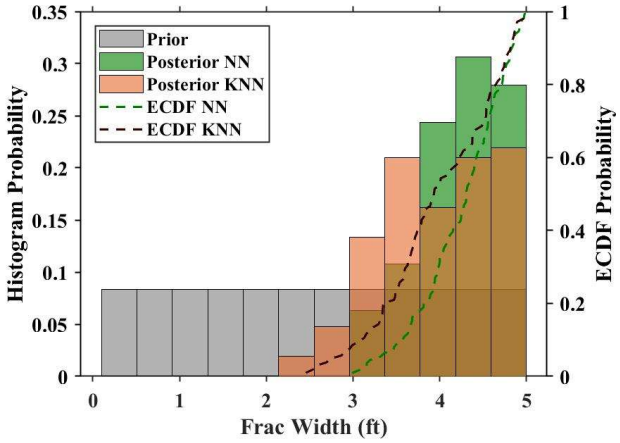
(e) Exponent of krg



(f) Endpoint of krw

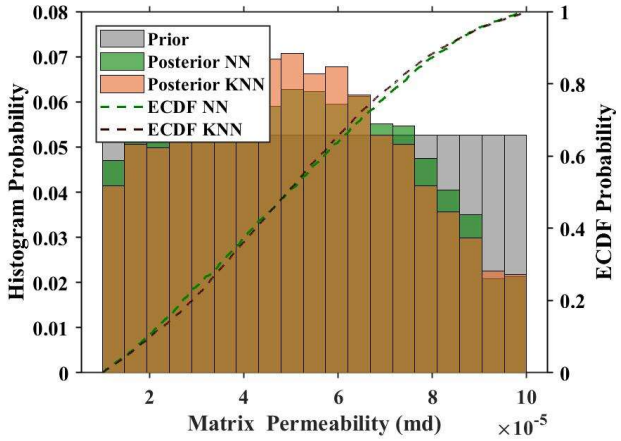


(g) Exponent of krw

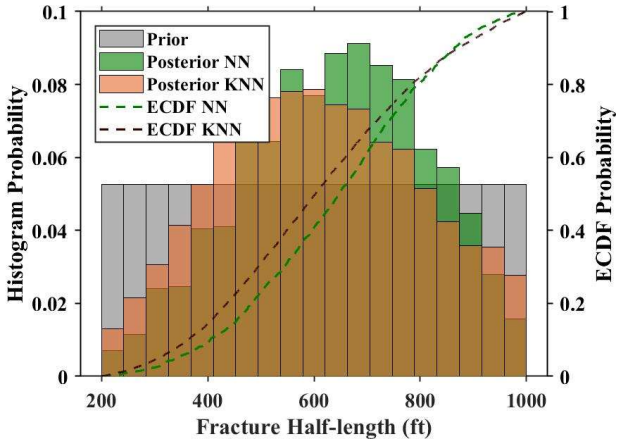


(h) Fracture width

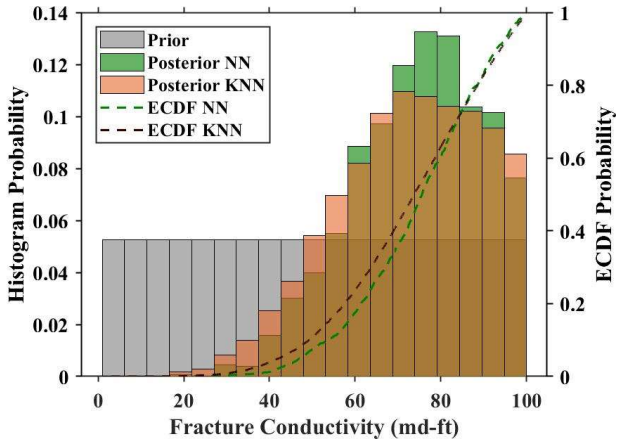
Figure 5.13: Comparison between prior and posterior distribution of uncertain parameter for NN (111 solutions) and KNN (105 solutions) proxy model : (a) Matrix permeability, (b) Fracture half-length, (c) Fracture conductivity, (d) Fracture water saturation, (e) Exponent of krg, (f) Endpoint of krw, (g) Exponent of krw and (h) Fracture width.



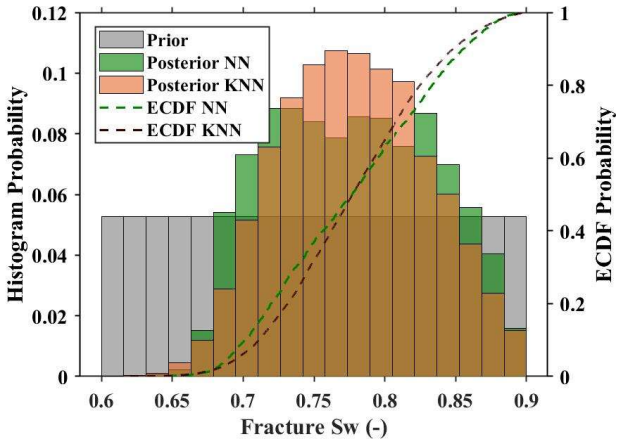
(a) Matrix permeability



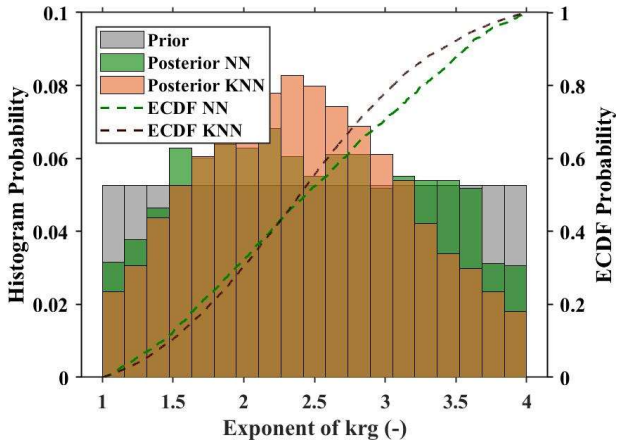
(b) Fracture half-length



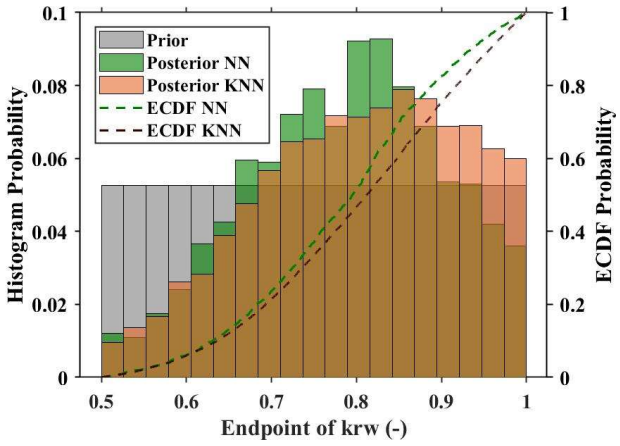
(c) Fracture conductivity



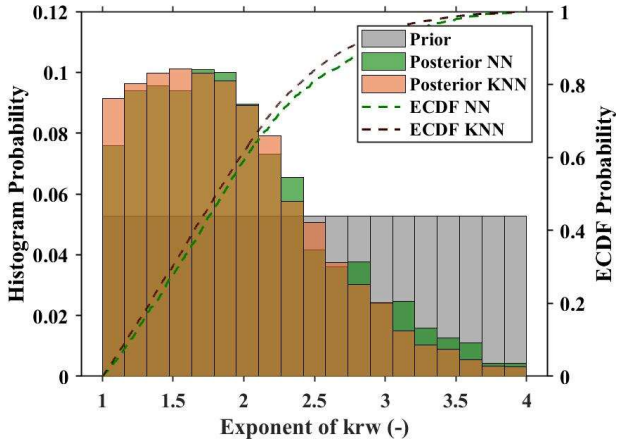
(d) Fracture water saturation



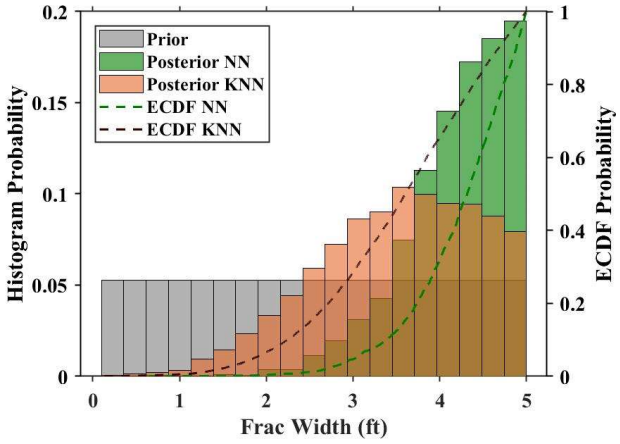
(e) Exponent of krg



(f) Endpoint of krw



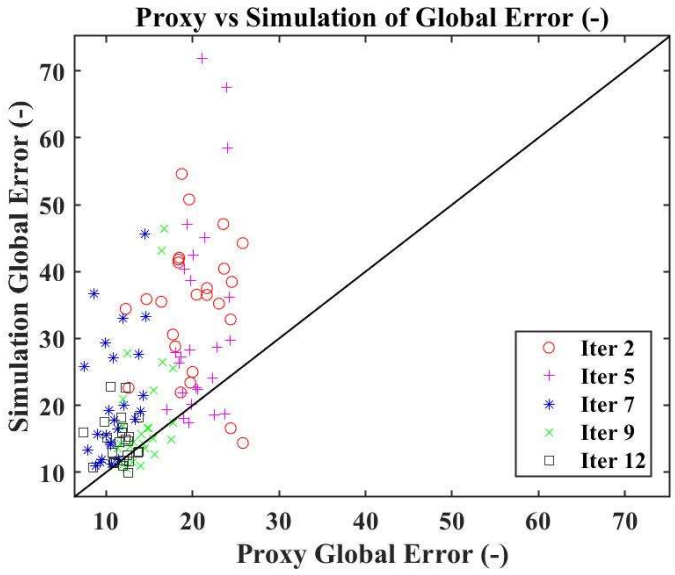
(g) Exponent of krw



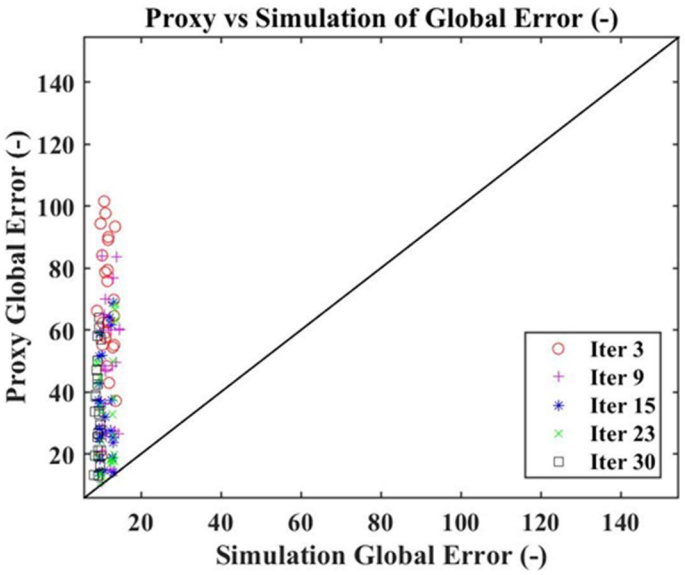
(h) Fracture width

Figure 5.14: Comparison between prior and posterior distribution of uncertain parameter for NN and KNN proxy model (1000 solutions from proxy): (a) Matrix permeability, (b) Fracture half-length, (c) Fracture conductivity, (d) Fracture water saturation, (e) Exponent of krg, (f) Endpoint of krw, (g) Exponent of krw and (h) Fracture width.

To compare the performance between NN and KNN, we plotted the proxy values vs simulation values cross plot for both NN and KNN as shown in Figure 5.15-5.17 for global error, BHP error and WGR error. The unit slope line represents 100% aligned between proxy and simulation values. As you can see that as iteration progressed, the proposed points came closer to unit slope line for both NN and KNN. However, NN predictability is obviously superior to KNN as more points located in the unit slope line in the later iteration while we can still observe the vertical lining up for KNN. The vertical line up implies that the proxy suggests the lower error values, but the simulation results values are still spreading over from low to high error values. This reflects a less accuracy and predictability of KNN than NN. All plots of 10 multiple proxies for both NN and KNN are also illustrated in Figure 5.18-5.19. Similarly, the proxy accuracy is improved as iteration progressed while NN provided more accuracy than KNN in all proxies and we still observed vertical lining up for KNN.

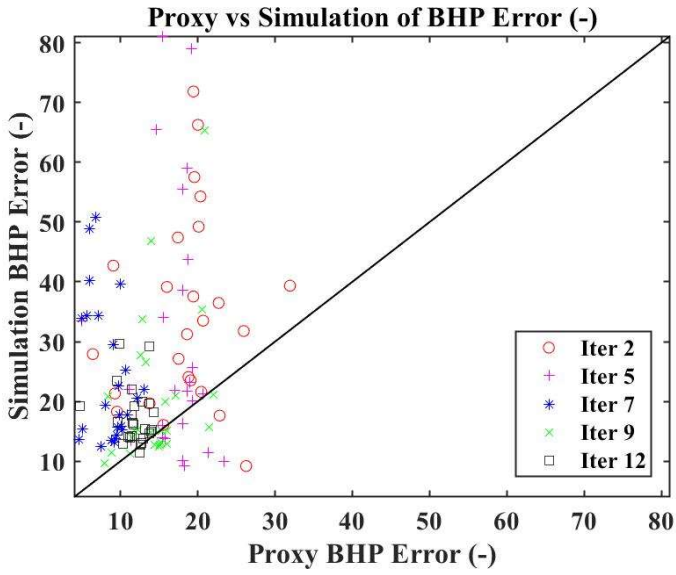


(a) Neural network

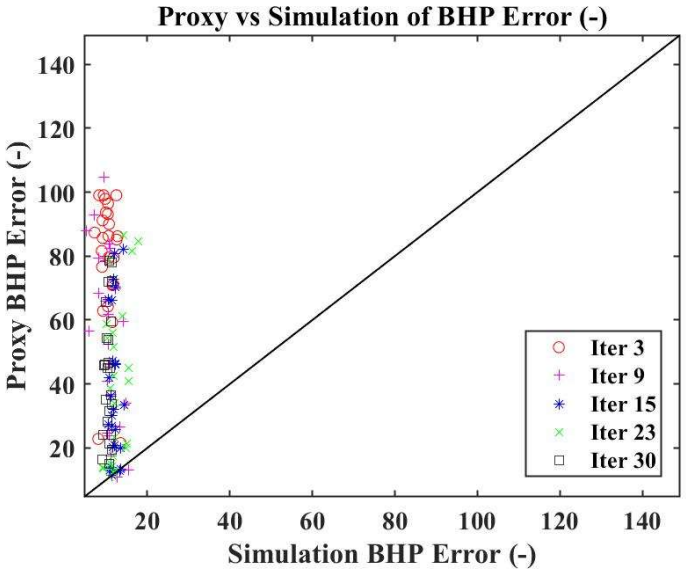


(b) K-nearest neighbors

Figure 5.15: Proxy vs. Simulation values of global error: (a) Neural network and (b) K-nearest neighbors.

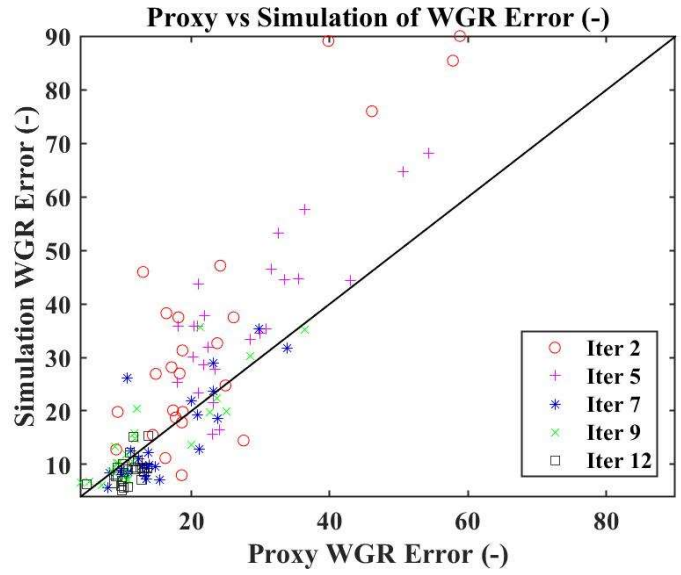


(a) Neural network

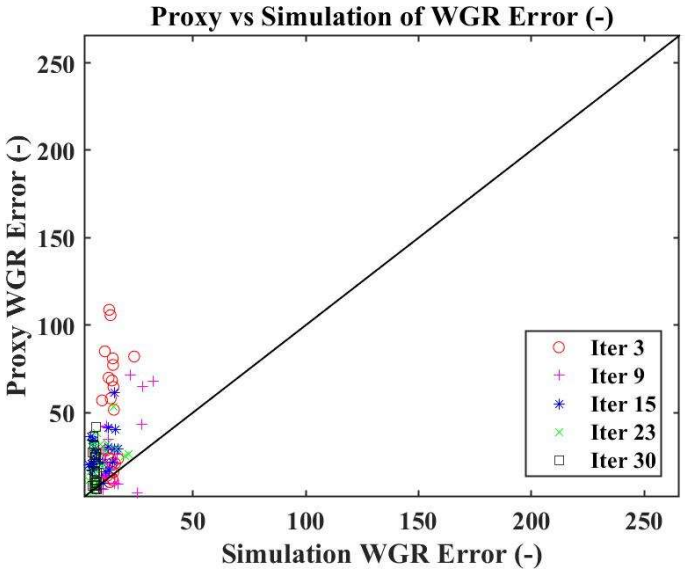


(b) K-nearest neighbors

Figure 5.16: Proxy vs. Simulation values of BHP error: (a) Neural network and (b) K-nearest neighbors.

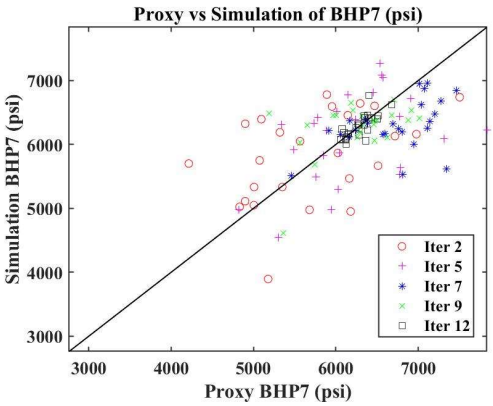


(a) Neural network

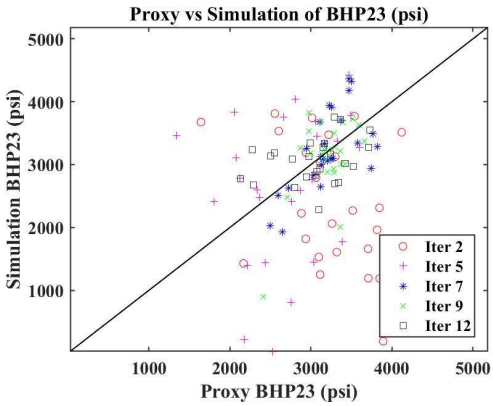


(b) K-nearest neighbors

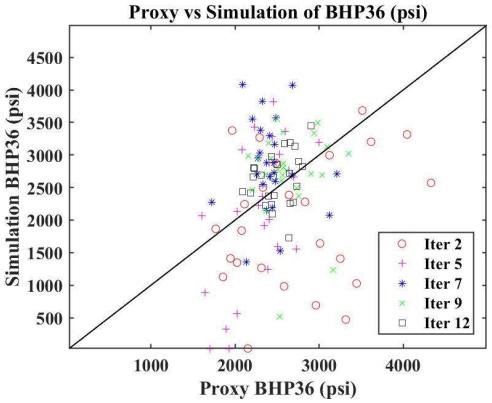
Figure 5.17: Proxy vs. Simulation values of WGR error: (a) Neural network and (b) K-nearest neighbors.



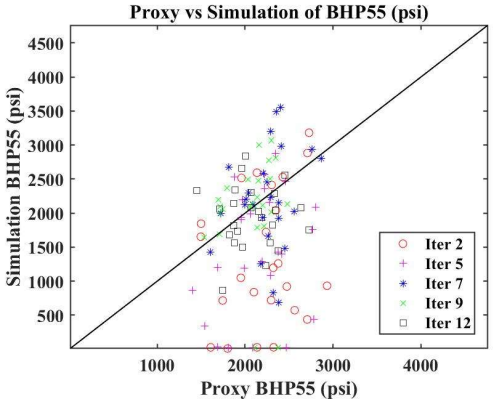
(a) BHP at day 7



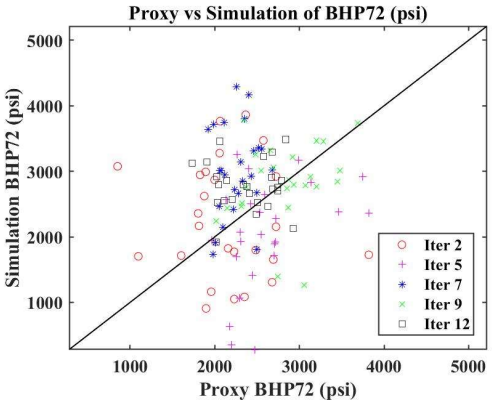
(b) BHP at day 23



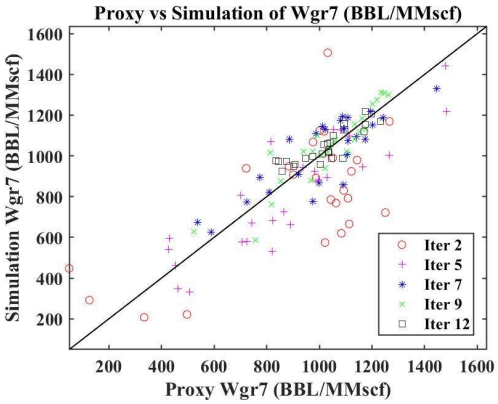
(c) BHP at day 36



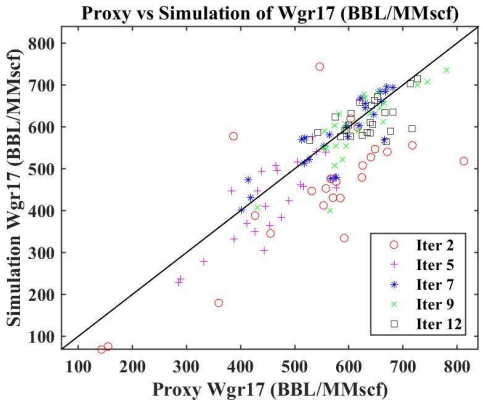
(d) BHP at day 55



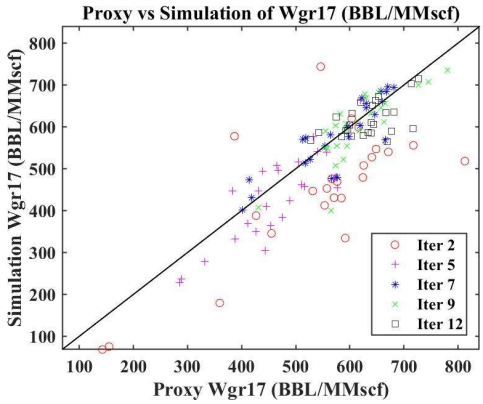
(e) BHP at day 72



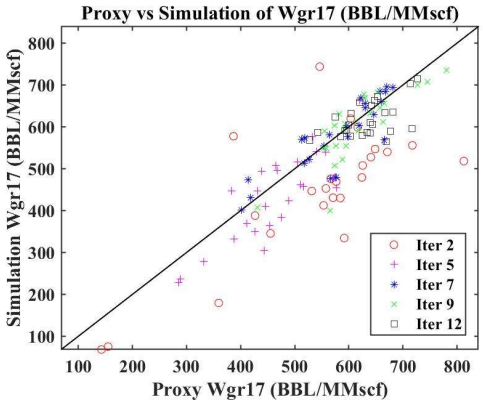
(f) WGR at day 7



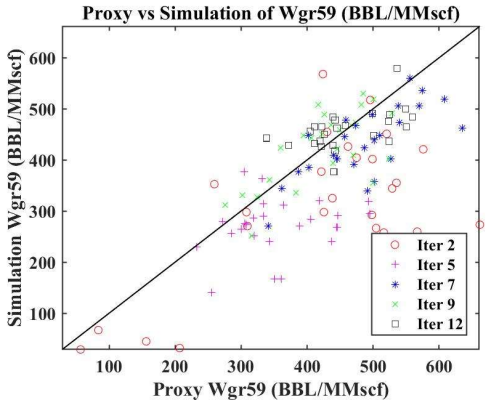
(g) WGR at day 17



(h) WGR at day 28

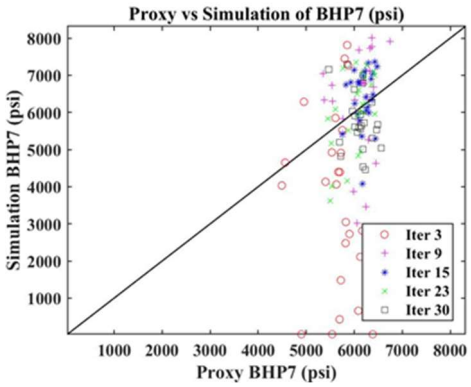


(i) WGR at day 40

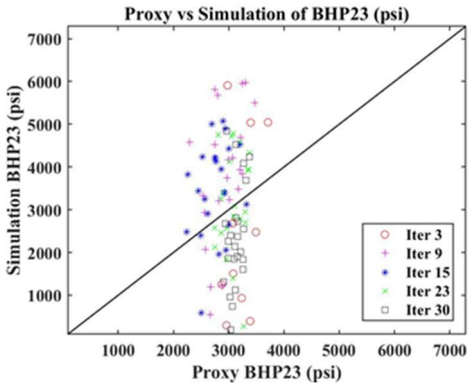


(j) WGR at day 59

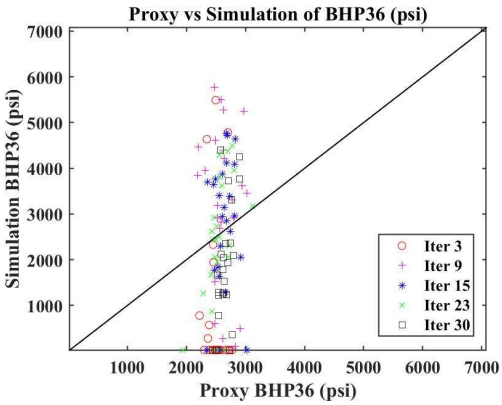
Figure 5.18: Cross plot between proxy and simulation values of 10 proxies for neural network as proxy model.



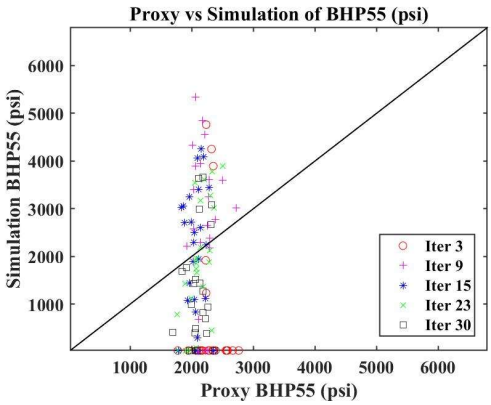
(a) BHP at day 7



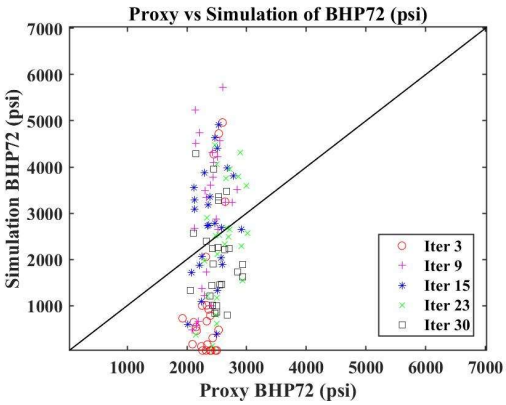
(b) BHP at day 23



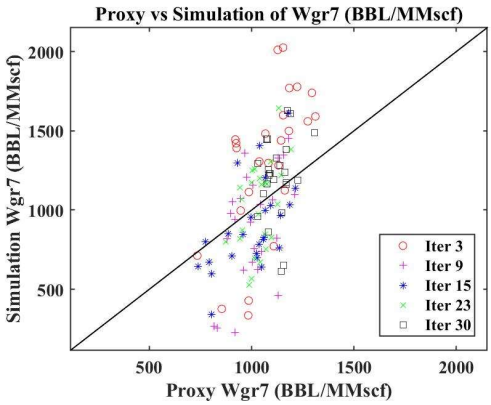
(c) BHP at day 36



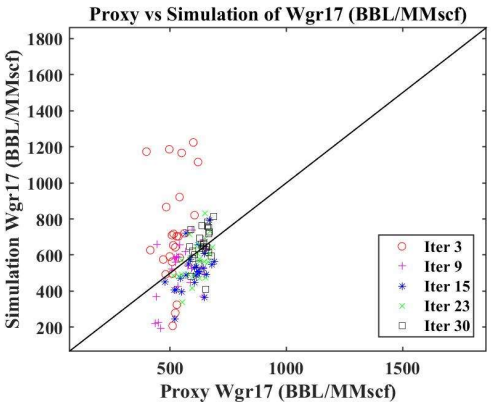
(d) BHP at day 55



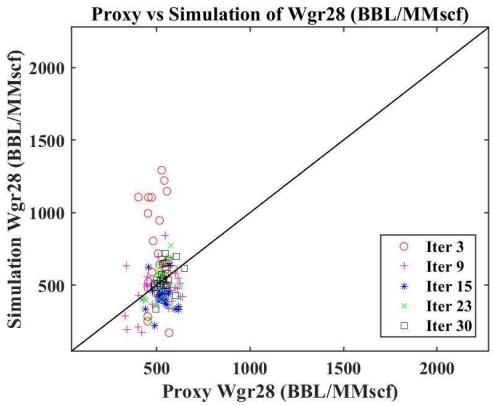
(e) BHP at day 72



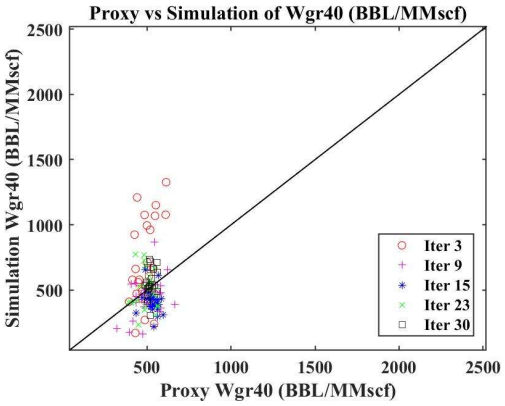
(f) WGR at day 7



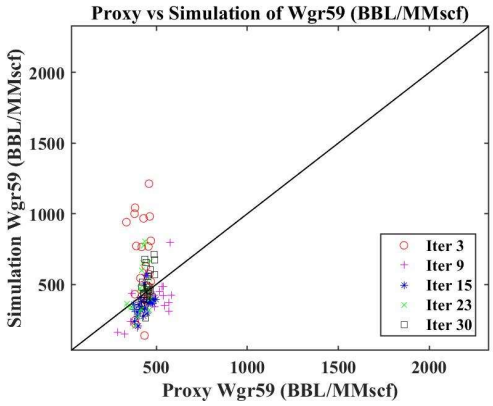
(g) WGR at day 17



(h) WGR at day 28



(i) WGR at day 40



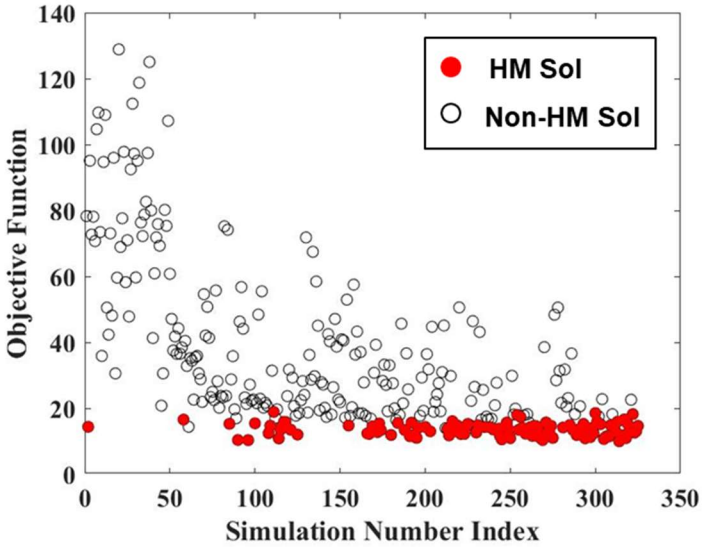
(j) WGR at day 59

Figure 5.19: Cross plot between proxy and simulation values of 10 proxies for k-nearest neighbors as proxy model.

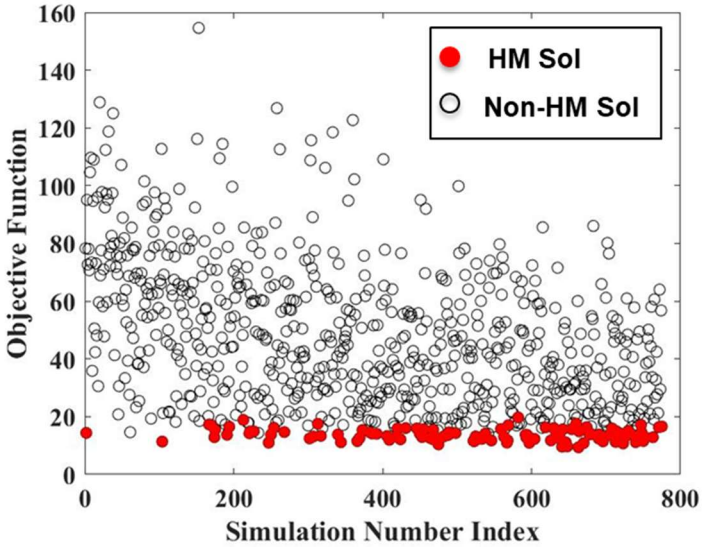
Figure 5.20 shows the objective function vs. simulation number index for both NN and KNN. This plot also emphasizes the better accuracy of NN over KNN as we observed faster improvement of objective function value. Figure 5.21-5.22 depicts scatter plots of the proposed cases in each iteration vs. simulation number index for each uncertain parameters for both NN and KNN, respectively. Firstly, the proxy-based MCMC algorithm proposed points over uncertain parameter domain; however, only some certain range of values can produce as history matching solutions. Secondly, we can see that the proposed

values were converged to the final range as simulation number index was higher for some parameters such as fracture half-length, fracture conductivity, fracture water saturation, exponent of krw and fracture width. This is because these parameters had more impact or more sensitive to the objection function. In sum, for both NN and KNN, Figure 5.20 reassures the accuracy and efficiency of the proxy-based MCMC algorithm for history matching.

For a total running time, we plotted the cumulative number of history matching solutions vs. elapsed time as shown in Figure 5.23. The elapsed time included all steps occurred in the AHM workflow such as automatic cases generating, proxy model training, proxy-based MCMC algorithm running, EDFM preprocessor running, simulation runs and stopping criteria evaluation. But the difference from using either NN or KNN would contribute to proxy model training, proxy-based MCMC algorithm and stopping criteria evaluation. By considering elapsed time required to obtain the same number of solutions, NN spent 1.6 times longer elapsed time than KNN. However, if we consider the number of simulation runs or total resources of simulation license as shown in Figure 5.24, NN performed significantly better than KNN, by obtaining 111 solutions from less total number of simulation runs than KNN by a factor of two. Therefore, in order to select which proxy model to be used, this would depend on each project limitation that either the running time or the available reservoir simulator resources are more important. In this study, we prioritize more on simulation runs resources; therefore, NN will be used to study in next section of the effect of EPA.

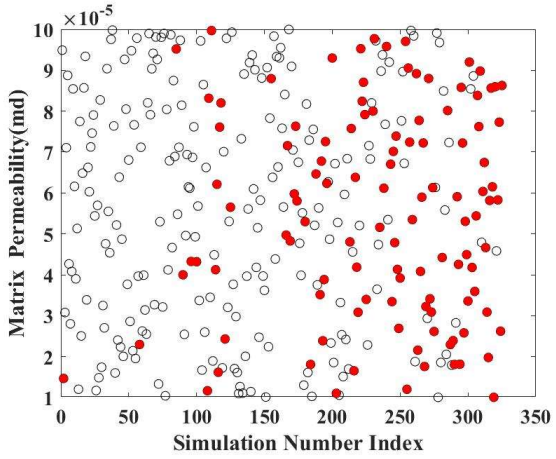


(a) Neural network

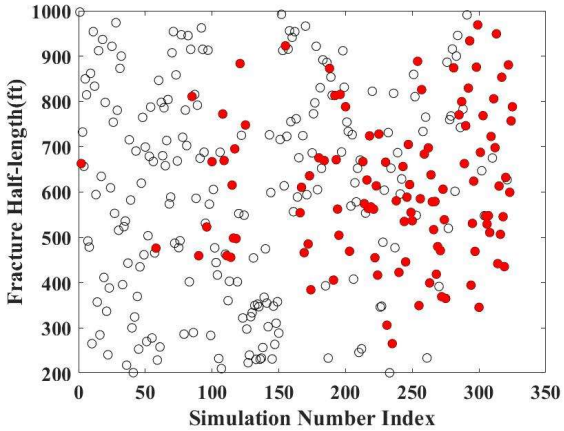


(b) K-nearest neighbors

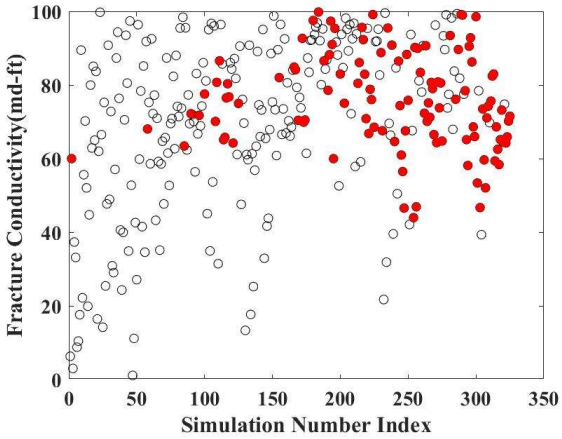
Figure 5.20: Scatter plot of objective function (global error) vs. simulation number index: (a) Neural network (b) K-nearest neighbors.



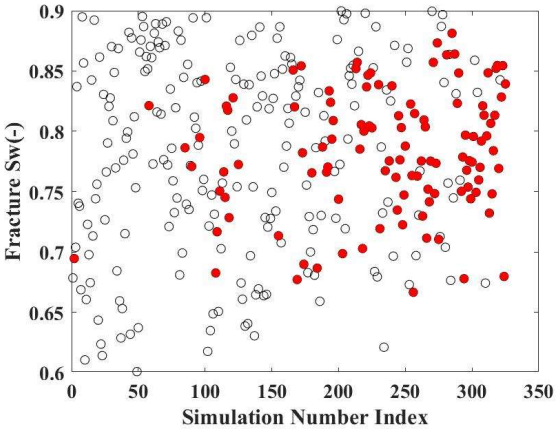
(a) Matrix permeability



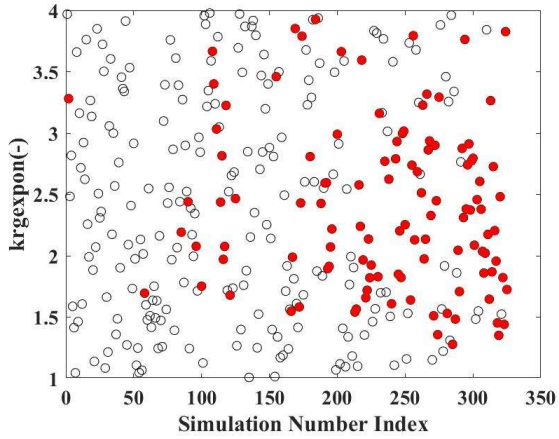
(b) Fracture half-length



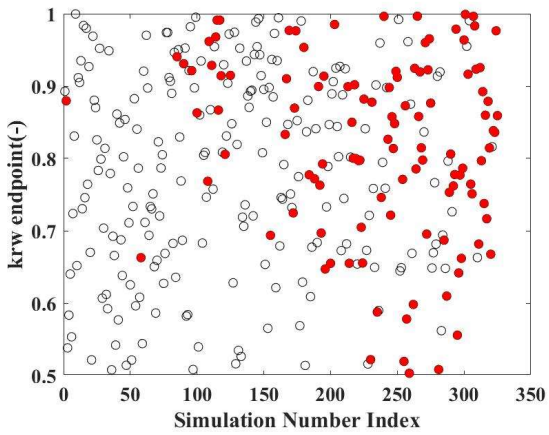
(c) Fracture conductivity



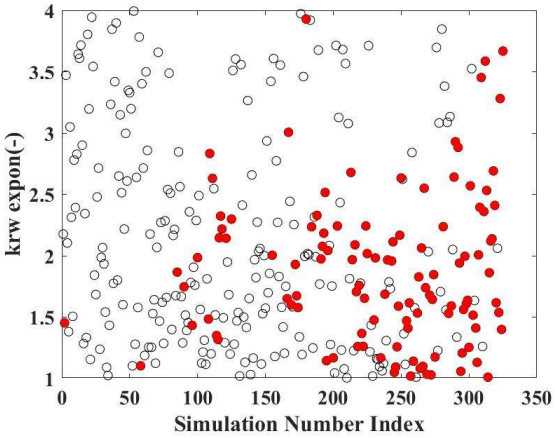
(d) Fracture water saturation



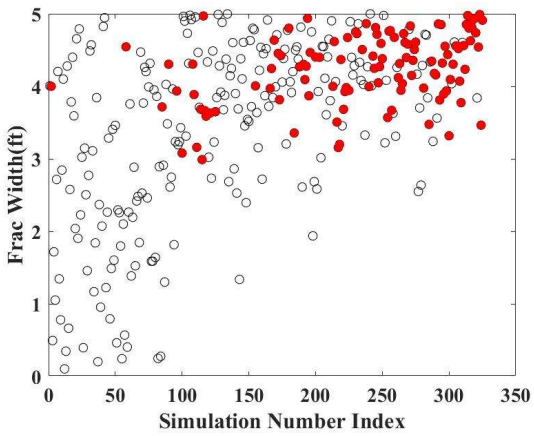
(e) Exponent of k_{rg}



(f) Endpoint of k_{rw}

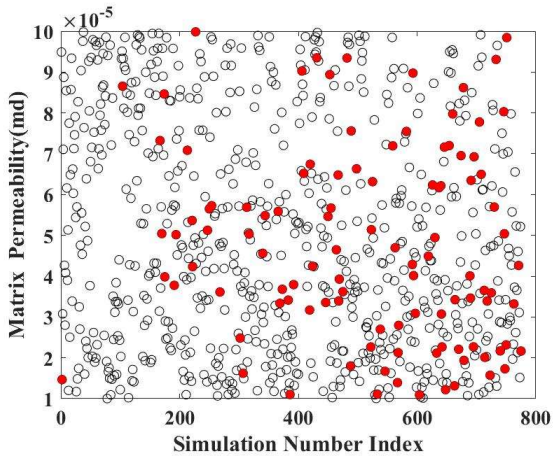


(g) Exponent of k_{rw}

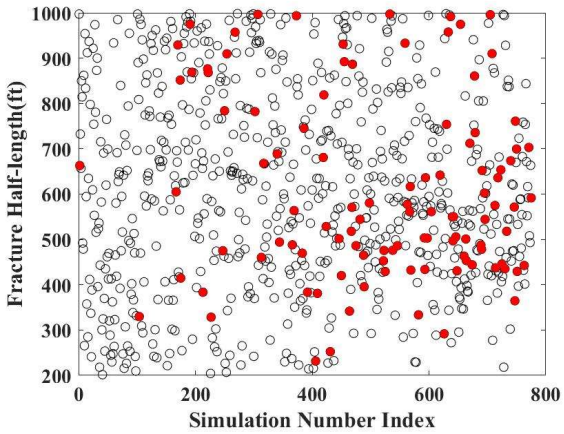


(h) Fracture width

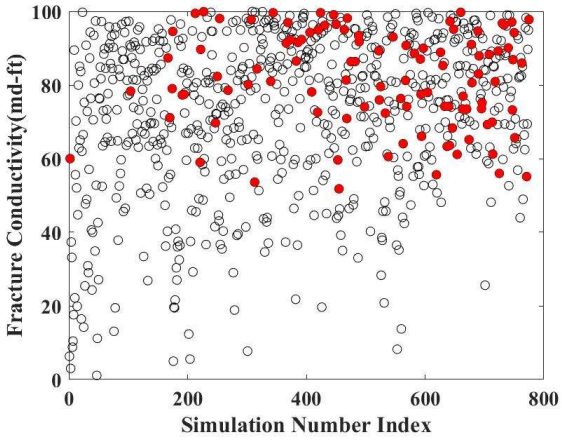
Figure 5.21: Uncertain parameter values vs. simulation number index for neural network. Black circles are non-HM solutions proposed by NN-MCMC algorithm in each iteration while filled red circles are HM solutions screened by the HM solution criteria.



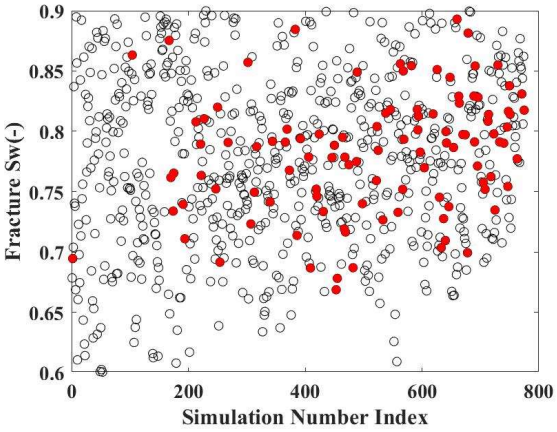
(a) Matrix permeability



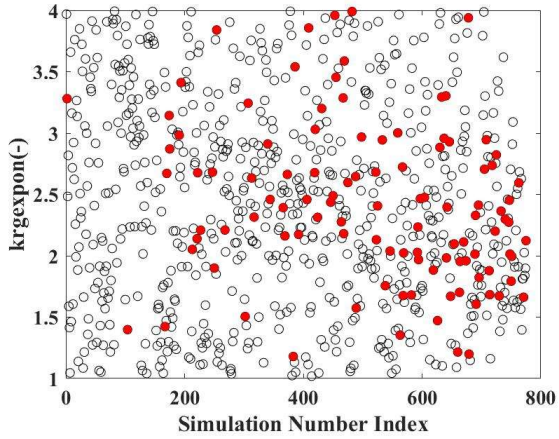
(b) Fracture half-length



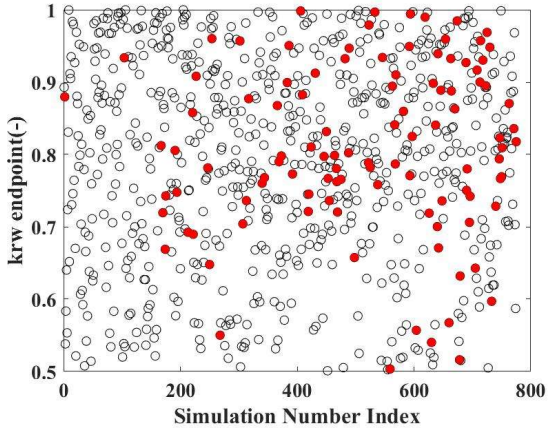
(c) Fracture conductivity



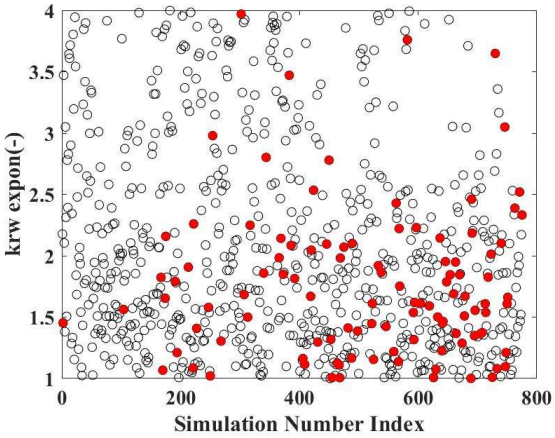
(a) (d) Fracture water saturation



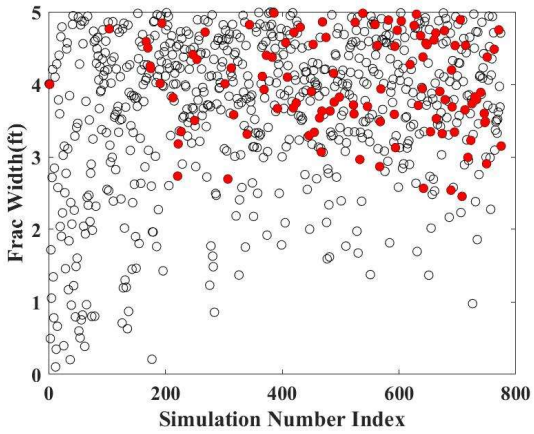
(e) Exponent of k_{rg}



(f) Endpoint of k_{rw}



(g) Exponent of k_{rw}



(h) Fracture width

Figure 5.22: Uncertain parameter values vs. simulation number index for k-nearest neighbors. Black circles are non-HM solutions proposed by NN-MCMC algorithm in each iteration while filled red circles are HM solutions screened by the HM solution criteria.

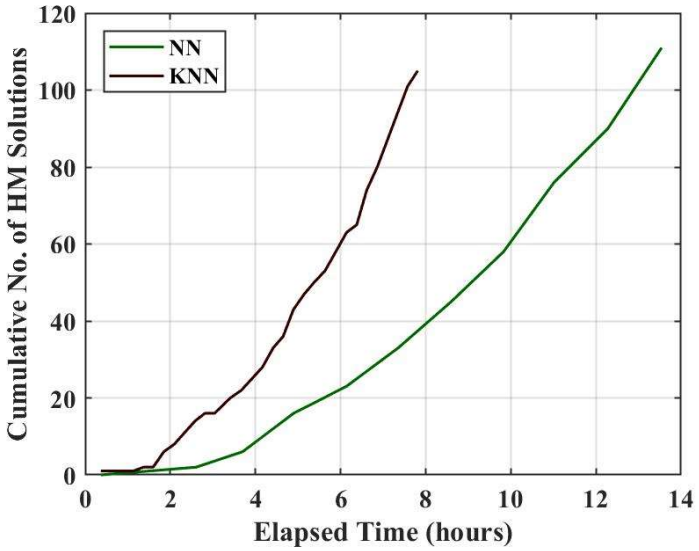


Figure 5.23: Comparison between NN and KNN for cumulative number of history matching solutions vs. elapsed time.

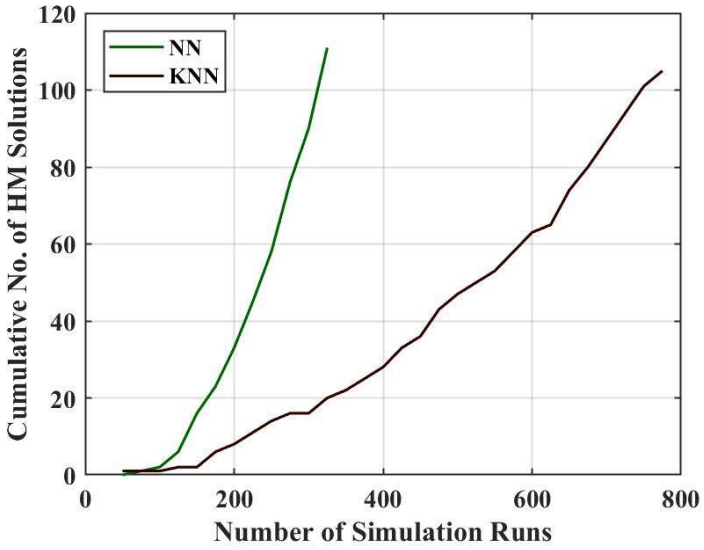


Figure 5.24: Comparison between NN and KNN for cumulative number of history matching solutions vs. number of simulation runs.

5.3.3 Probabilistic Forecasting

Following the multiple history matching solutions that we obtained from the AHM workflow, we can perform probabilistic production forecast for all solutions of both NN and KNN. In this study, we forecasted production for 30 years by constraining BHP at 500 psi, as shown in Figure 5.25. The production forecast of cumulative gas production for both NN and KNN is shown in Figure 5.26 with P10-50-90 cases. After that, we can estimate the empirical cumulative distribution function (ECDF) of gas estimated ultimate recovery (EUR) from all cases as shown Figure 5.27. The figures illustrated both the solutions from the AHM workflow and 1000 solutions generated by the proxy models. Both ECDF from AHM workflow and proxy models are quite in line to each other but more difference is seen in KNN due to its nature of lower predictability as the proxy model. To compare between NN and KNN, the P10-50-90 of forecasted gas EUR are in the same range with a slight discrepancy of 10%.

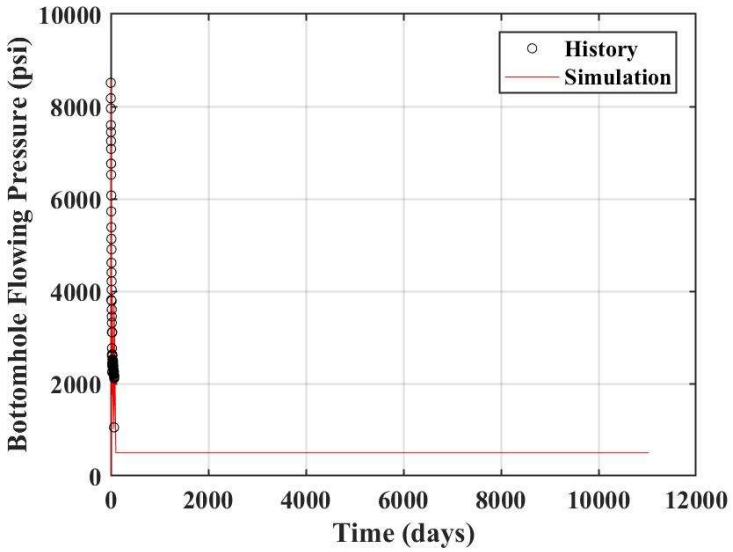
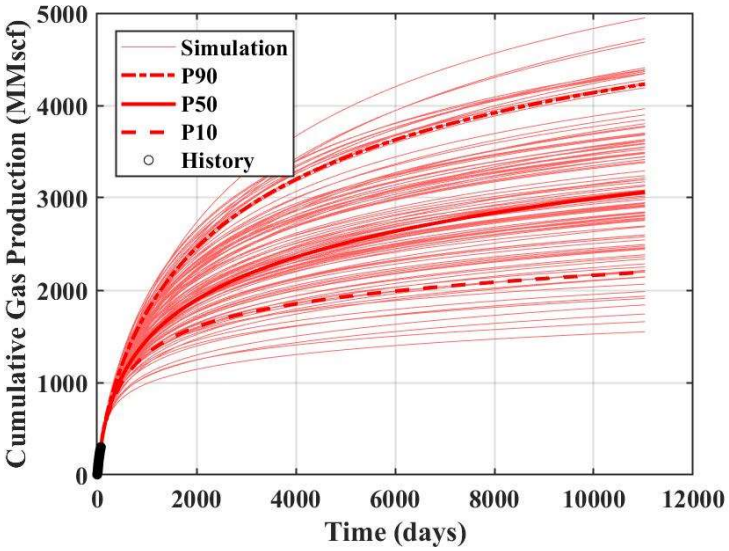
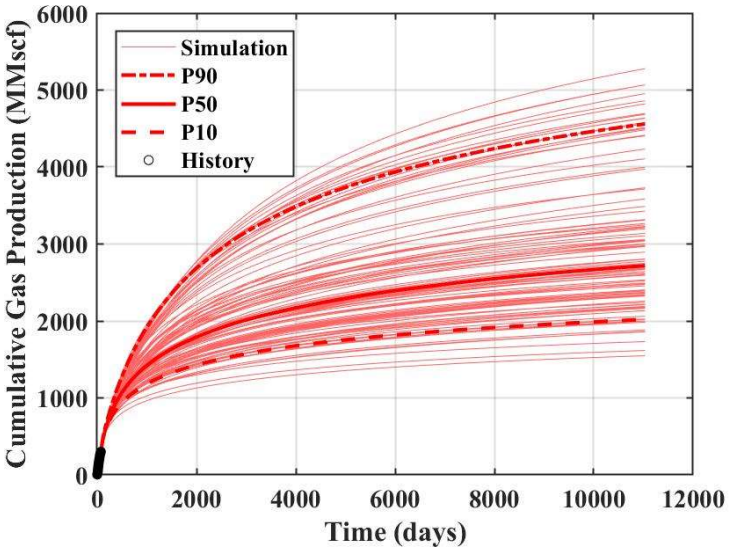


Figure 5.25: BHP of 500 psi was used as production constraint for 30-year production forecast.

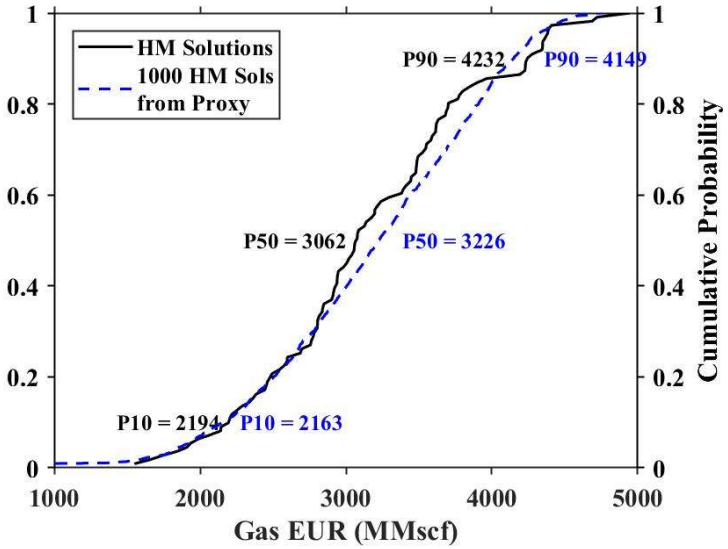


(a) Neural network

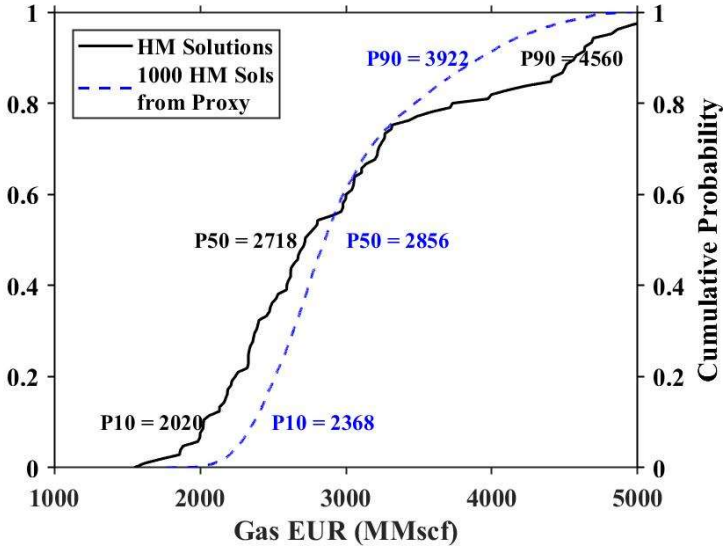


(b) K-nearest neighbors

Figure 5.26: Cumulative gas production production forecast of all history matching solutions from AHM workflow: (a) Neural network and (b) K-nearest neighbors.



(a) Neural network



(b) K-nearest neighbors

Figure 5.27: Empirical cumulative distribution function (ECDF) of gas estimated ultimate recovery (EUR) for history matching solutions from AHM workflow and 1000 history matching solutions generated by proxy model: (a) Neural network and (b) K-nearest neighbors.

5.4 Comparison of Case with Enhanced Permeability Area (EPA)

In this section, we performed another scenario to investigate how the EPA existence around main hydraulic fractures would impact the history matching solutions and production forecast. We used NN as proxy for proxy-based MCMC algorithm.

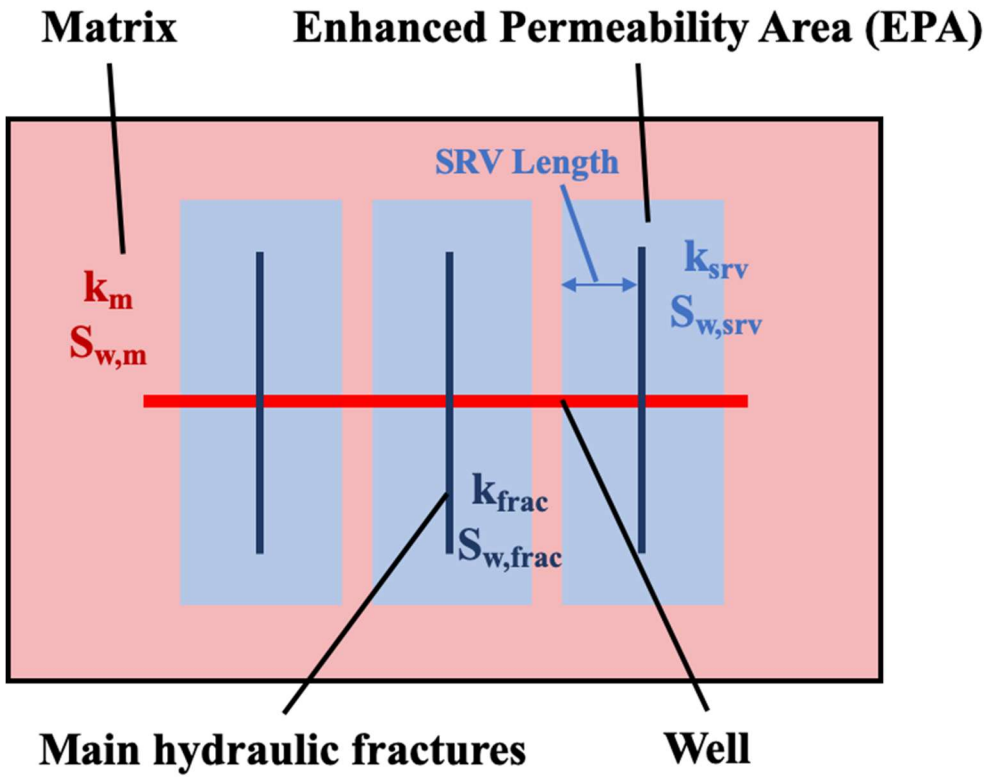


Figure 5.28: The concept of enhanced permeability area (EPA) with main hydraulic fractures.

5.4.1 Parameters Identification & Screening and History Matching

For parameter identification and screening, we used the screening results from section 5.3 that all determined parameters will be used further in history matching. To model EPA, we identified two additional parameters including Simulated Rock Volume

(SRV) permeability or the permeability in the EPA and the length of SRV from hydraulic fractures as illustrated in Figure 5.28. In this study, we assumed initial range of SRV permeabilities to be around 10% of hydraulic fracture conductivity range, and SRV length is from no EPA area to fully connected across fractures spacing. For SRV length, besides we modified permeability as SRV permeability, we also modified water saturation inside the SRV length the same as fracture water saturation. All parameters ranges for EPA case are summarized in Table 5.4.

Code	Uncertain parameter	Unit	Distribution	Min value	Max value
A	Matrix permeability	mD	Uniform	0.00001	0.0001
B	Fracture half-length	ft	Uniform	200	1000
C	Fracture conductivity	md-ft	Uniform	1	100
D	Fracture water saturation	-	Uniform	0.6	0.9
E	Exponent of k_{ro}	-	Uniform	1	4
F	Endpoint of k_{rw}	-	Uniform	0.5	1
G	Exponent of k_{rw}	-	Uniform	1	4
H	Fracture width	ft	Uniform	0.01	5
I	SRV permeability	md	Uniform	0.1	1
J	SRV length	ft	Uniform	0	27.5

Table 5.4: Summary of ten uncertain parameters and their prior distributions for EPA case

The AHM workflow continued as previously explained in section 5.3. In this case, the workflow also stopped at the maximum iteration of 12. 72 history-matching solutions

were found from total 325 simulation runs or around 22% comparing to 34% for the case with only hydraulic fractures (without EPA). The lower of number solutions found is obviously from a higher number of uncertain parameters including additional two EPA parameters. The number of history matching solutions vs. iteration number is shown in Figure 5.29. The simulation results of all 72 solutions are shown in Figure 5.30. In addition, the best match cases for both without EPA (section 5.3) and with EPA are shown in Figure 5.31 and 5.32, respectively, and Table 5.5 summarizes uncertain properties of the best cases for both scenario

.

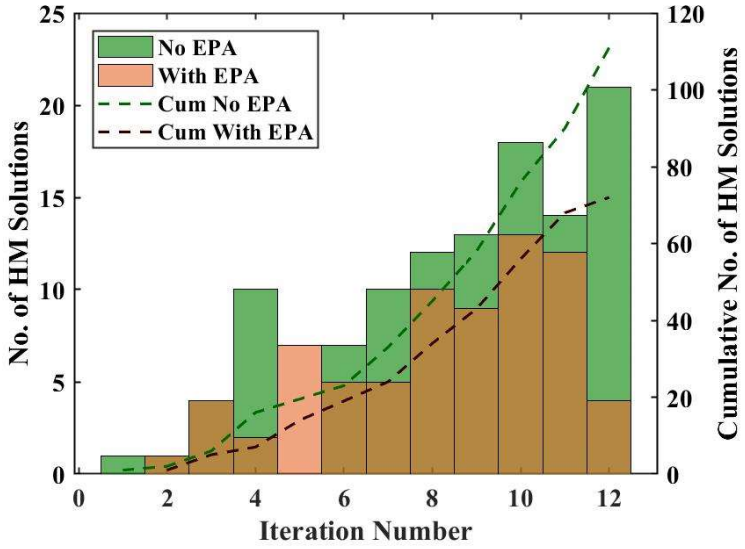
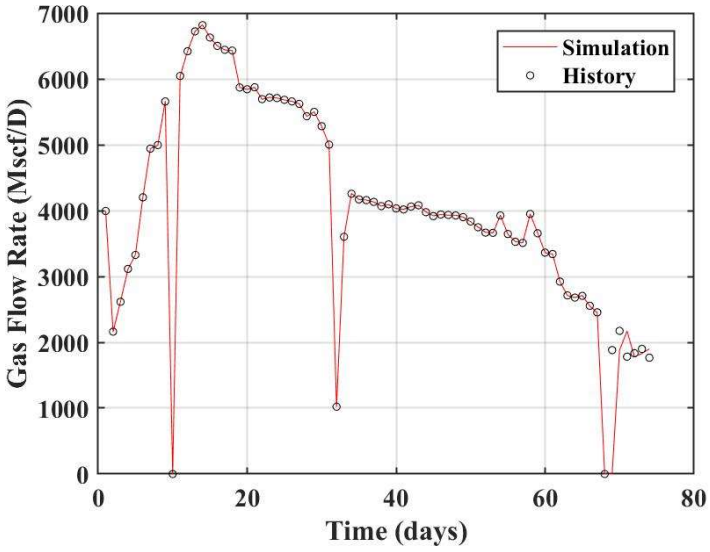
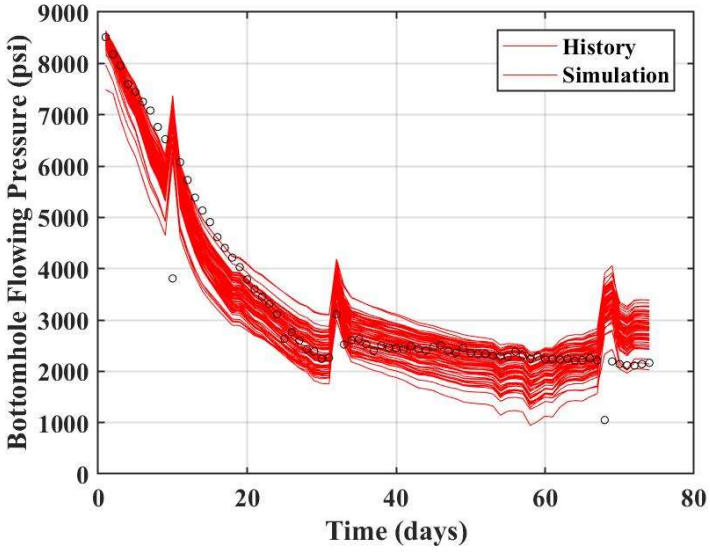


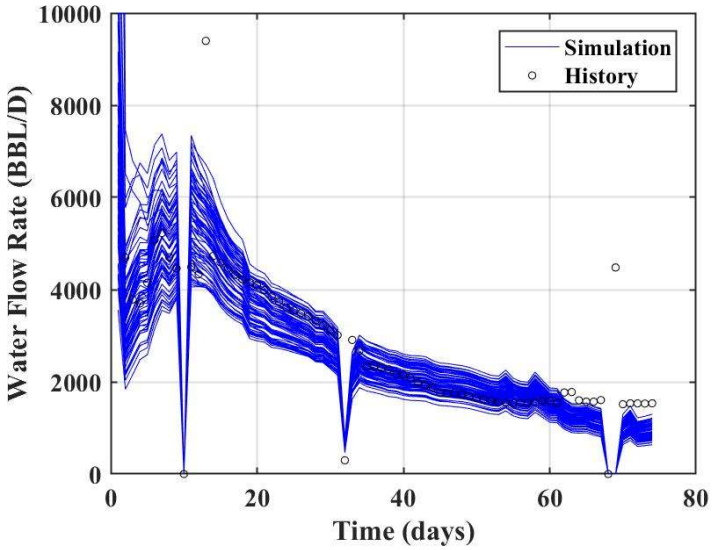
Figure 5.29: Comparison of number of history matching solutions vs. iteration number between cases without EPA and with EPA.



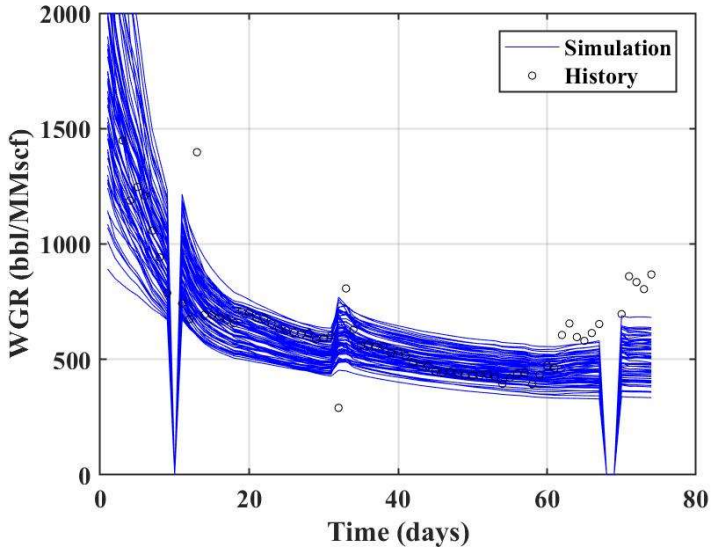
(a) Gas flow rate as constraint



(b) Flowing BHP

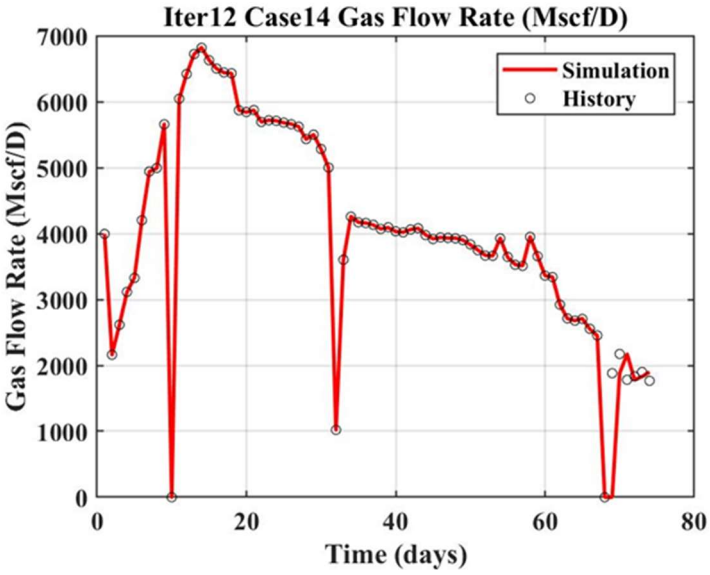


(c) Water flow rate

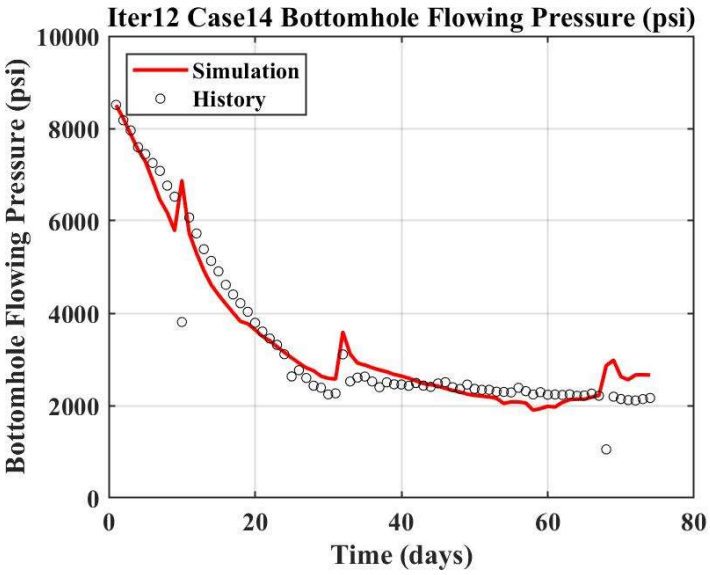


(d) WGR

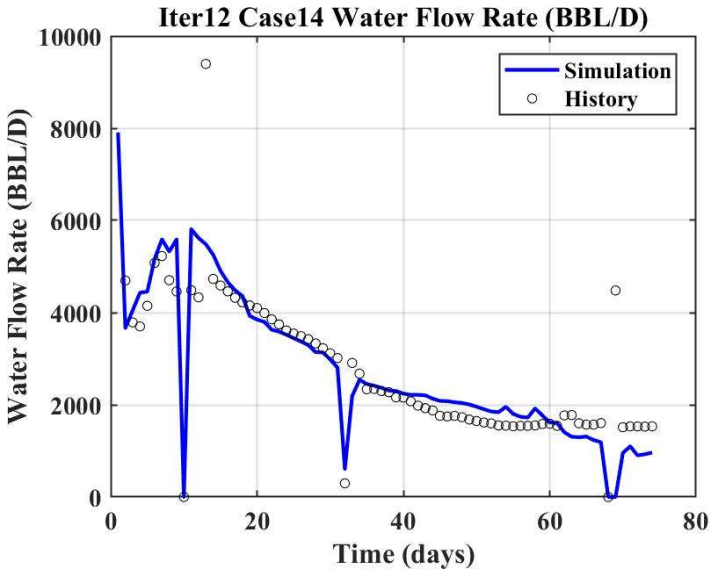
Figure 5.30: Comparison between simulation results and historical data of 72 history matching solutions for case with EPA: (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) WGR.



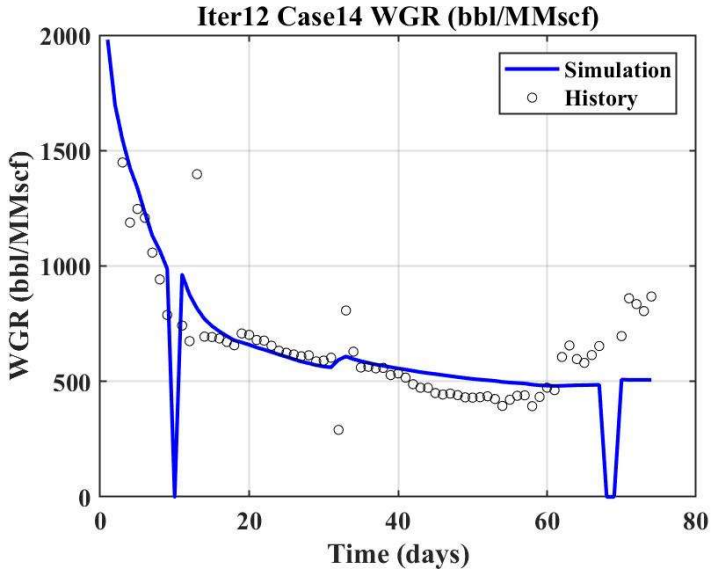
(a) Gas flow rate as constraint



(b) Flowing BHP

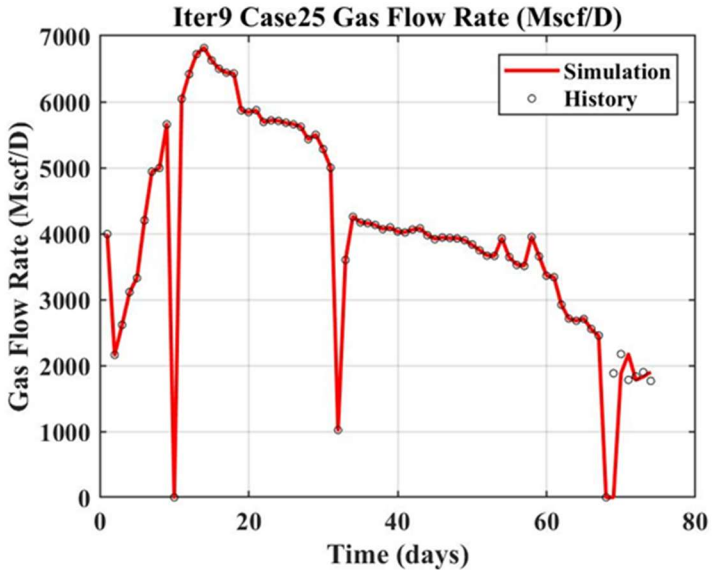


(c) Water flow rate

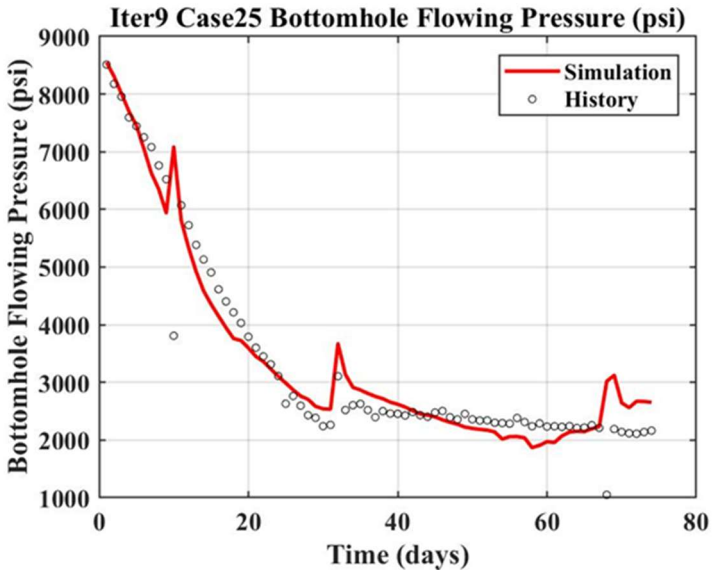


(d) WGR

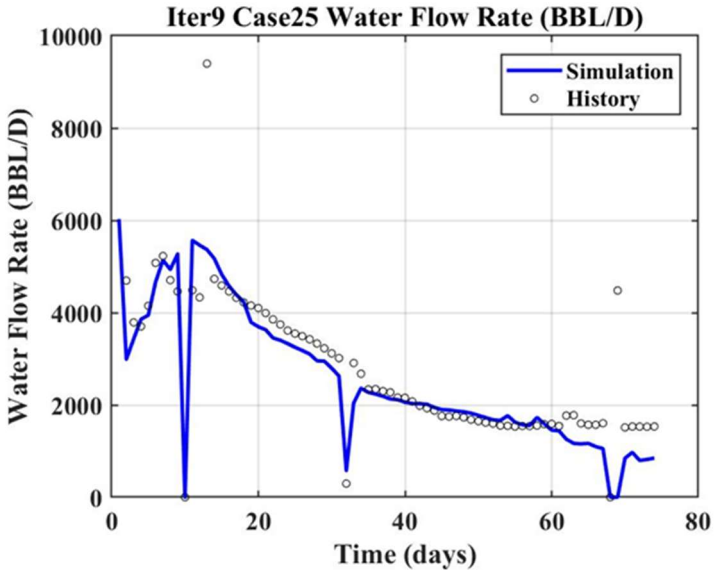
Figure 5.31: Comparison between simulation results and historical data of the best match case (iteration 12 case 14) for hydraulic fracture only or without EPA (section 5.3): (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) WGR.



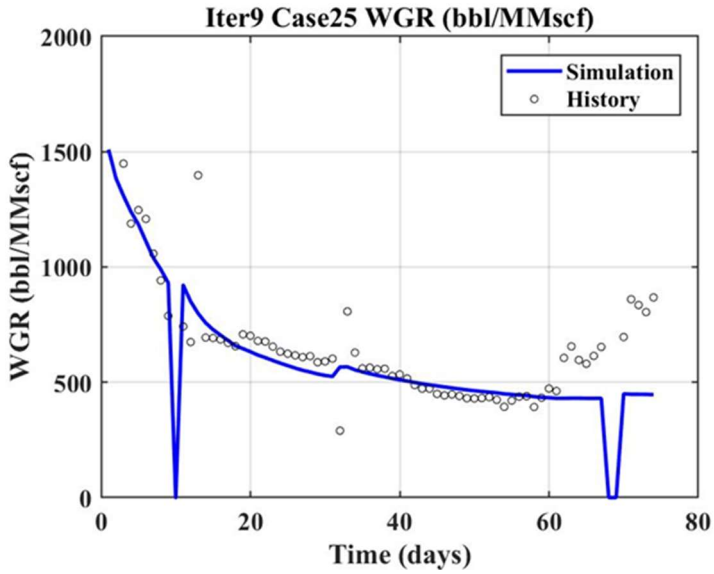
(a) Gas flow rate as constraint



(b) Flowing BHP



(c) Water flow rate



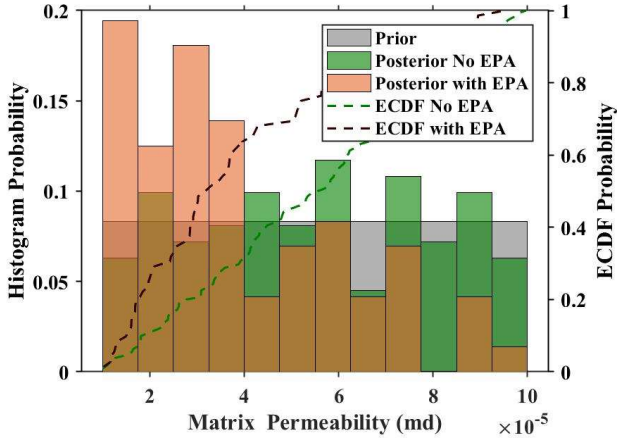
(d) WGR

Figure 5.32: Comparison between simulation results and historical data of the best match case for the case with EPA: (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) WGR.

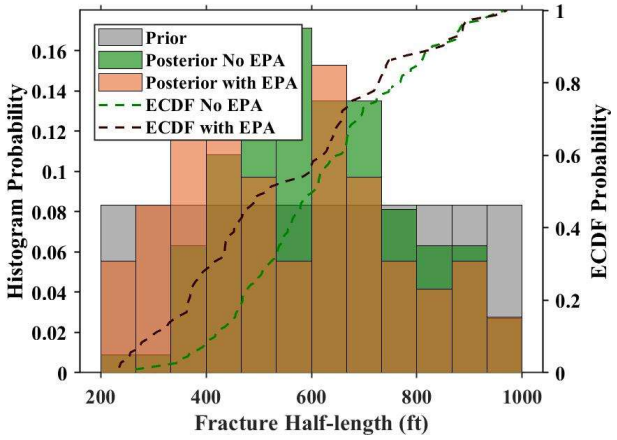
Uncertain parameter	Unit	Without EPA (Iteration 12 Case14)	With EPA (Iteration 9 Case 25)
Matrix permeability	mD	0.000031	0.0000024
Fracture half-length	ft	442	379
Fracture conductivity	md-ft	59	53
Fracture water saturation	-	0.81	0.81
Exponent of k_{ro}	-	1.87	1.71
Endpoint of k_{rw}	-	0.89	0.92
Exponent of k_{rw}	-	1.01	1.68
Fracture width	ft	4.98	3.79
SRV permeability	md	n/a	0.55
SRV length	ft	n/a	18.4

Table 5.5: Summary of ten uncertain parameters of the best match cases for both without EPA and with EPA.

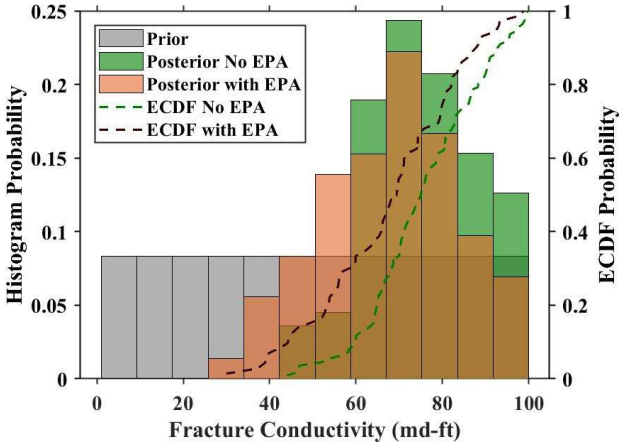
Similarly, we obtained posterior distribution of uncertain parameters including the additional uncertain parameters for EPA including SRV permeability and SRV length. The comparison of each uncertain parameter between case without EPA and with EPA is illustrated in Figure 5.33. We can see that when there is EPA around the main hydraulic fractures, matrix permeability, fracture half-length, fracture conductivity, fracture water saturation and fracture width are shifted to lower values. This is because there are more contributed areas to the fluid flow by the EPA around main hydraulic fractures.



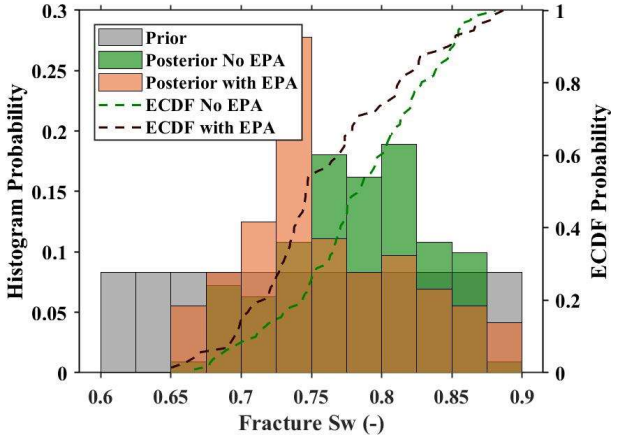
(a) Matrix permeability



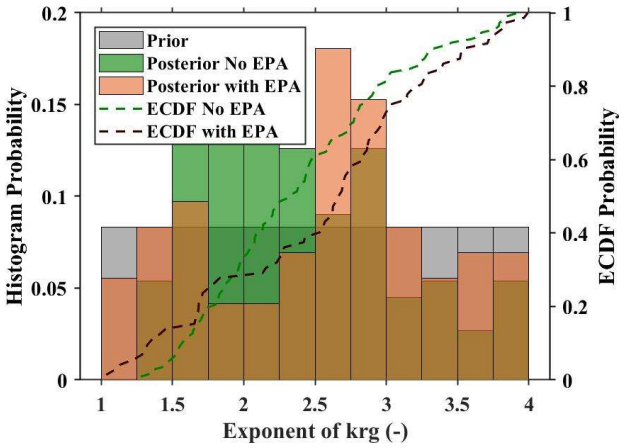
(b) Fracture half-length



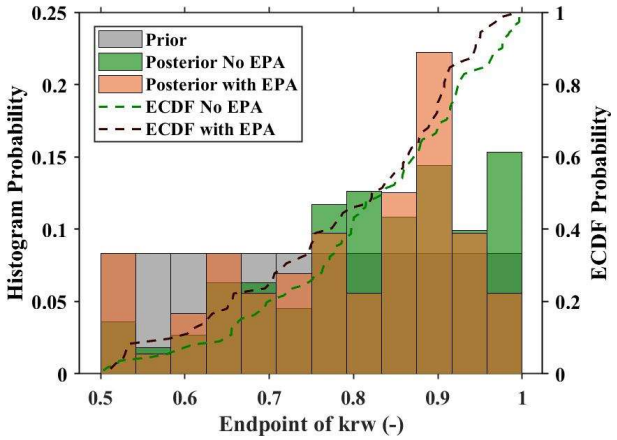
(c) Fracture conductivity



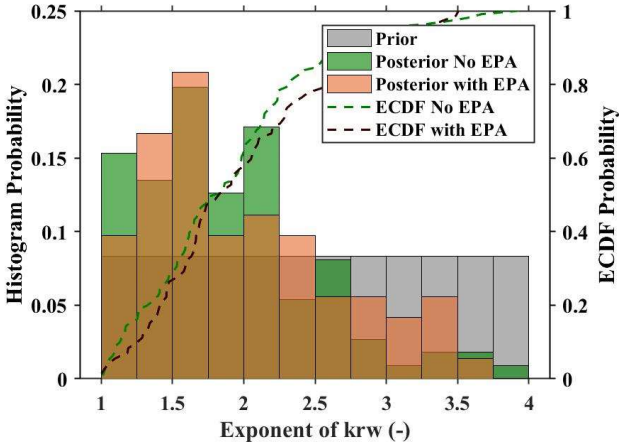
(d) Fracture water saturation



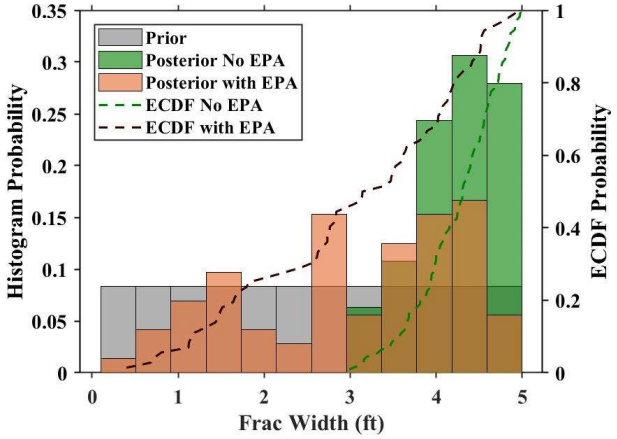
(e) Exponent of krg



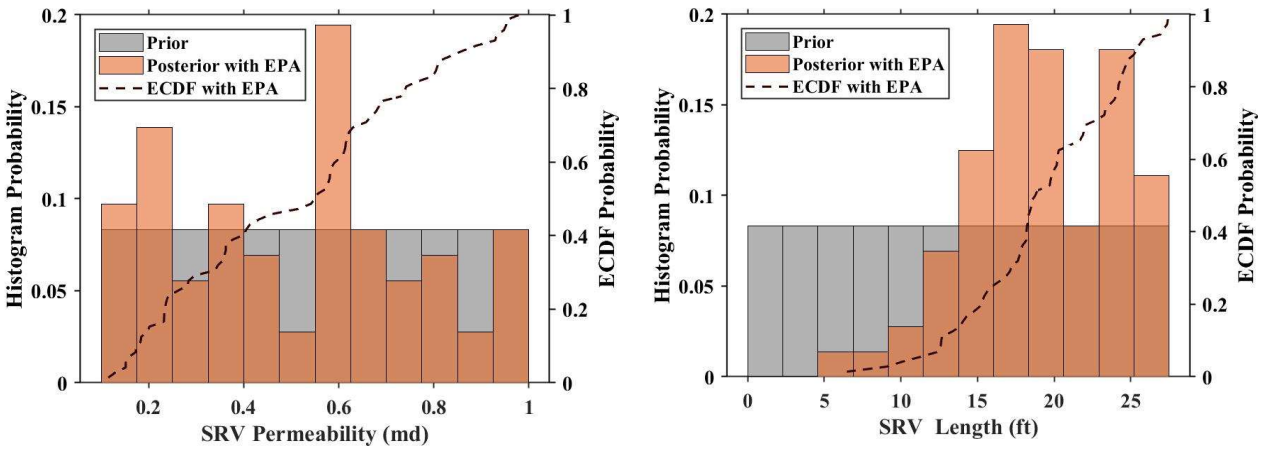
(f) Endpoint of krw



(g) Exponent of krw



(h) Fracture width

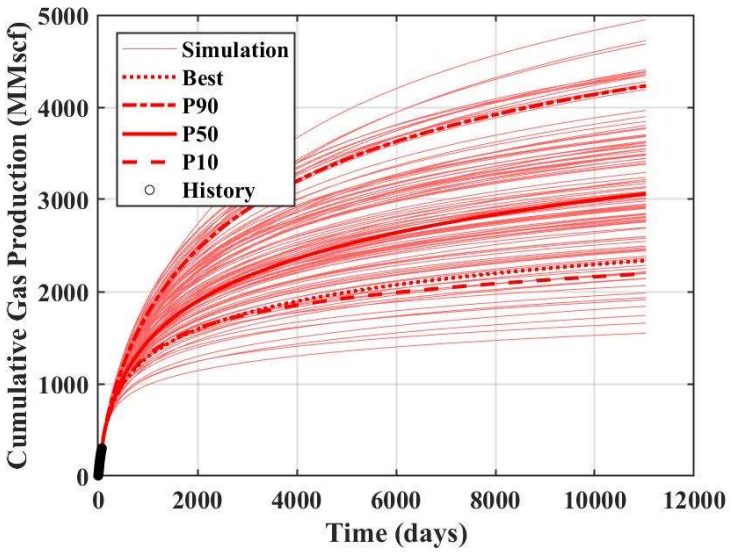


(i) SRV permeability (j) SRV length

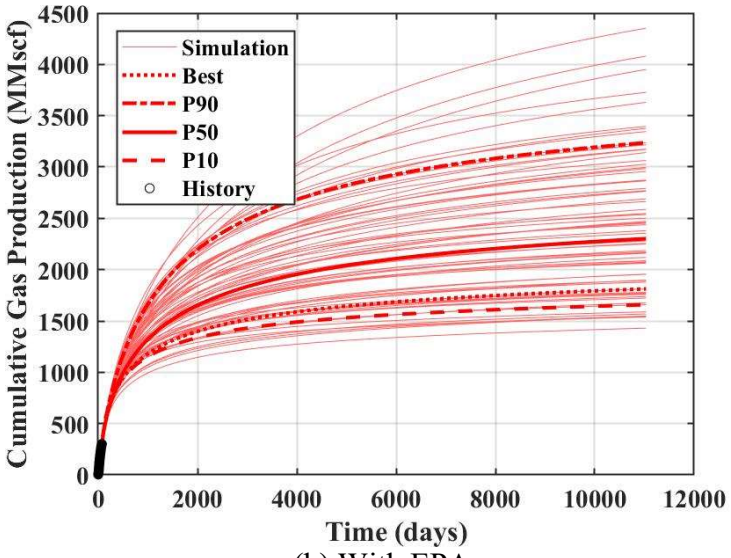
Figure 5.33: Comparison between prior and posterior distribution of uncertain parameter for case without EPA (111 solutions) and case with EPA (72 solutions): (a) Matrix permeability, (b) Fracture half-length, (c) Fracture conductivity, (d) Fracture water saturation, (e) Exponent of krg, (f) Endpoint of krw, (g) Exponent of krw, (h) Fracture width, (i) SRV permeability and (j) SRV length.

5.4.2 Probabilistic Forecasting

After we obtained history-matching solutions, we can perform production forecast in the probabilistic manner as previously explained in section 5.3.3. The predicted gas productions for the case without EPA and with EPA of all 72 solutions are shown in Figure 5.34a and 5.34b, respectively. Also, the Gas EUR ECDF for the case without EPA and with EPA are illustrated in Figure 5.35a and 5.35b respectively. We can observe that gas EUR is predicted around 25% lower than the case without EPA in section 5.3, comparing between P50 of 3062 MMscf and 2302 MMScf for the case without EPA and with EPA, respectively. This is because the lower fracture geometry cases were obtained from history matching when EPA existed around the hydraulic fractures.

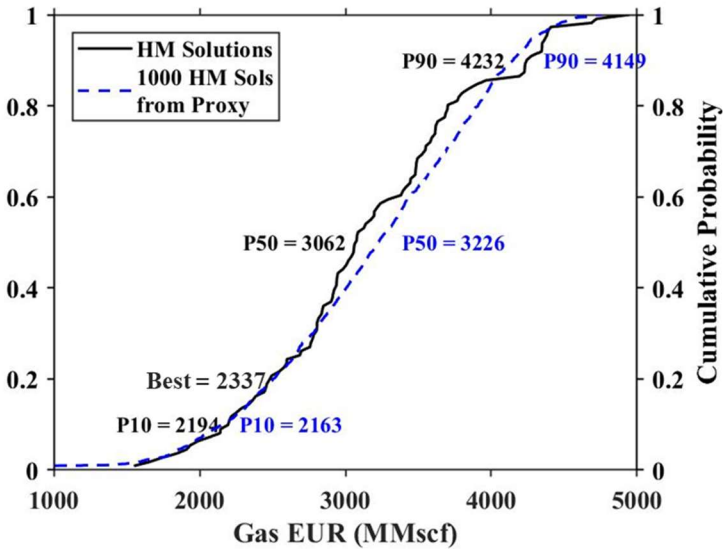


(a) Hydraulic fractures only (Without EPA)

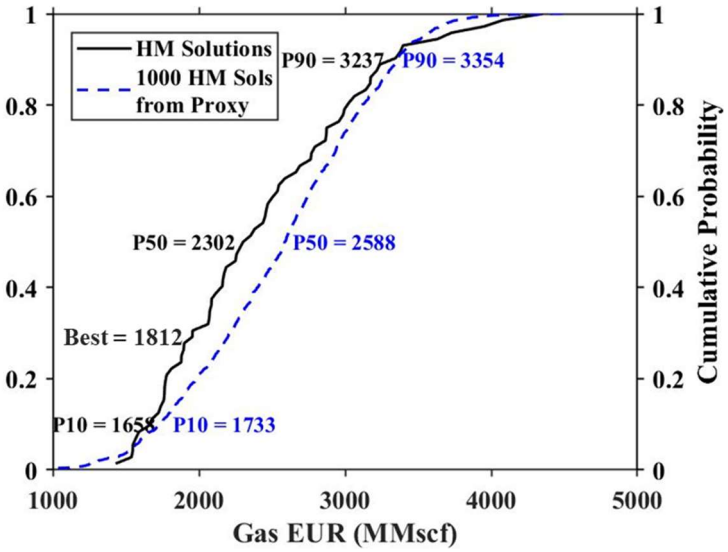


(b) With EPA

Figure 5.34: Cumulative gas production forecast of all 72 history matching solutions from AHM workflow with best match case and P10-50-90 cases: (a) Hydraulic fractures only (Without EPA) (b) With EPA.



(a) Hydraulic fractures only (Without EPA)



(b) With EPA

Figure 5.35: Empirical cumulative distribution function (ECDF) of gas estimated ultimate recovery (EUR) for 72 history matching solutions and 1000 history matching solutions generated by proxy model (neural network): (a) Hydraulic fractures only (Without EPA) (b) With EPA.

To sum up, when the EPA existed around hydraulic fractures, the history matching solutions changed to lower values of hydraulic fracture geometries and lower matrix permeability. These changes triggered the discrepancy of production forecast between the case with and without EPA. The lower gas EUR production forecast was obtained for the case with EPA. This means that if we assume a simple model with only hydraulic fractures but in a realistic situation, there are EPA due to the small fracture networks created around main hydraulic fractures, we would overpredict the fracture geometry and Gas EUR prediction, in this case around 25-30%.

The pressure distributions at different times from best match cases of two scenarios, without EPA and with EPA, are illustrated in Figure 5.36. We can see that pressure between each hydraulic fracture cluster for the case without EPA is higher than the case with EPA at each time step. However, the fracture half-length for case without EPA is longer than the case with EPA. Therefore, this causes the case without EPA to drain more gas as illustrated by a wider area of depleted pressure than the case with EPA, and certainly, resulting in a higher gas EUR as discussed in Figure 5.35.

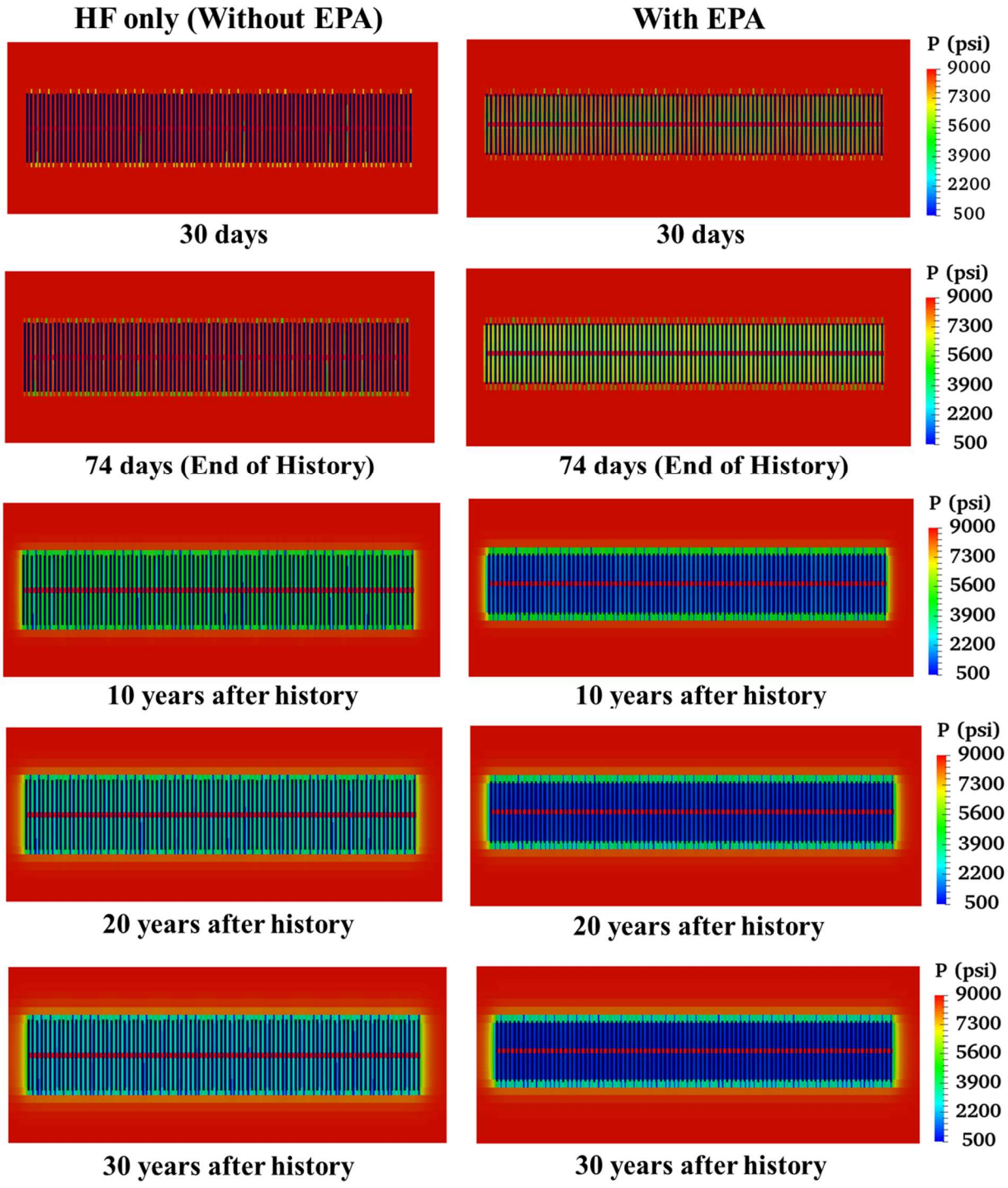


Figure 5.36: Pressure distributions comparison between case with hydraulic fractures only (without EPA) and case with hydraulic fractures and EPA (with EPA).

5.4 CONCLUSIONS

We applied the AHM workflow to shale gas well and the performances of different proxy models such as NN and KNN were firstly investigated. NN performs better than KNN in terms of predictive accuracy when validating with the actual simulator. In other words, NN requires a smaller number of simulation runs than KNN in order to obtain the same number of solutions. On the other hand, KNN performs better in terms of running time, as it required almost half of running time compared with NN. We also investigated the impact of EPA existence to the history matching solutions and production forecast. With EPA, the lower fracture geometry solutions were obtained in the history matching solutions compared with the case of hydraulic fractures only (without EPA). This causes the production forecast of the case with EPA to be significantly lower than the one with only hydraulic fractures (without EPA). This means that if a simple model with only hydraulic fractures was assumed while in the actual operation, there is EPA due to the small fracture networks created around main hydraulic fractures, we would overpredict the fracture geometry and gas EUR prediction.

Chapter 6: Application of the Assisted History Matching Workflow to Shale Gas Well with Natural Fractures using Neural-Network Markov Chain Monte Carlo Algorithm

6.1 INTRODUCTION

The objective of this chapter is to emphasize the application of the Assisted History Matching (AHM) workflow to a shale gas well. We used proxy-based MCMC algorithm and neural network (NN) as proxy model for history matching. In addition, we investigated the effect of natural fractures (NF) existence on the posterior distribution of history matching solutions and production forecast. We performed two cases including hydraulic fractures only (No NF) and hydraulic fractures and natural fractures (With NF).

6.2 RESERVOIR MODEL

For this study, we constructed a homogeneous reservoir model due to the fact that the static model integrating geological and petrophysical inputs were not available. A model dimension is 6600 ft × 2950 ft in x and y direction and a model dimension in z direction is equal to fracture height. The grid block dimension is 50 ft × 50 ft × fracture height. The initial reservoir pressure is 7397 psi. The reservoir model description is summarized in Table 6.1. Also, the well and fracture information are summarized in Table 6.2. The horizontal well has a lateral length of 5800 ft. As shown in Figure 6.1, the well was completed with 20 stages of hydraulic fractures with a total cluster number of 58. For the fractures modeling, we used EDFM method to handle fractures non-intrusively to the main reservoir grid.

For shale gas reservoirs, we included the gas desorption effect using Langmuir isotherm as shown in Table 6.3. Also, we considered the pressure-dependent permeability for both matrix and fractures due to a geomechanical effect during reservoir depletion. The

permeability reduction happens when reservoir pressure was depleted. As shown in Equation 5.1, the permeability will be decreased when reservoir is depleted with the permeability modulus of 0.055 and 0.056 for matrix and fractures, respectively. The normalized pressure-dependent permeability curves for both matrix and fracture are illustrated in Figure 6.2.

Reservoir description	**Value**	**Unit**
Model dimension (x × y)	6000 × 2950	ft
Number of grid blocks (x × y × z)	120 × 59 × 1	-
Initial reservoir pressure	7397	psi
Reservoir temperature	200	^{o}F
Residual water saturation	20%	-
Total compressibility	3×10-6	psi^{-1}
Matrix porosity	6%	-

Table 6.1: Summary of basic reservoir properties used in this study

Well and fractures description	**Value**	**Unit**
True vertical depth of a lateral section	8228	ft
Well length	5800	ft
Number of stages	20	-
Number of clusters	58	-
Cluster spacing	100	ft

Table 6.2: Summary of well and hydraulic fractures information

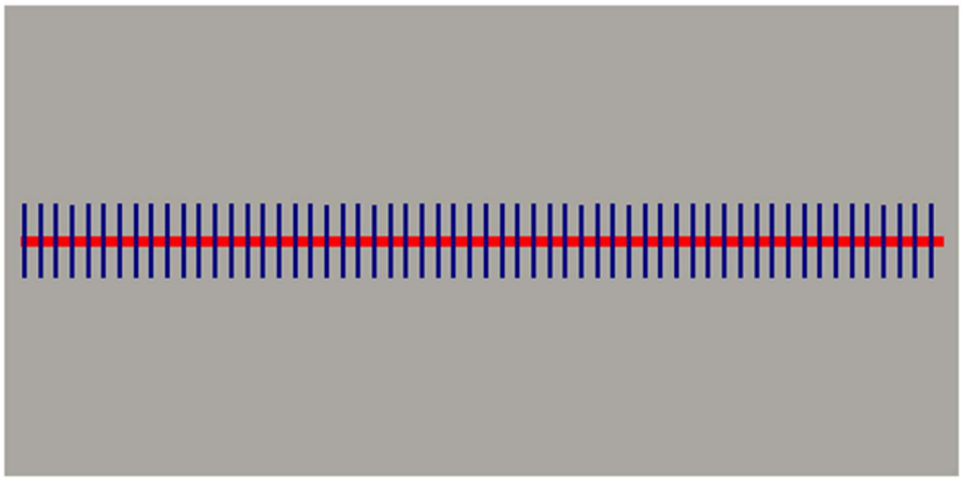
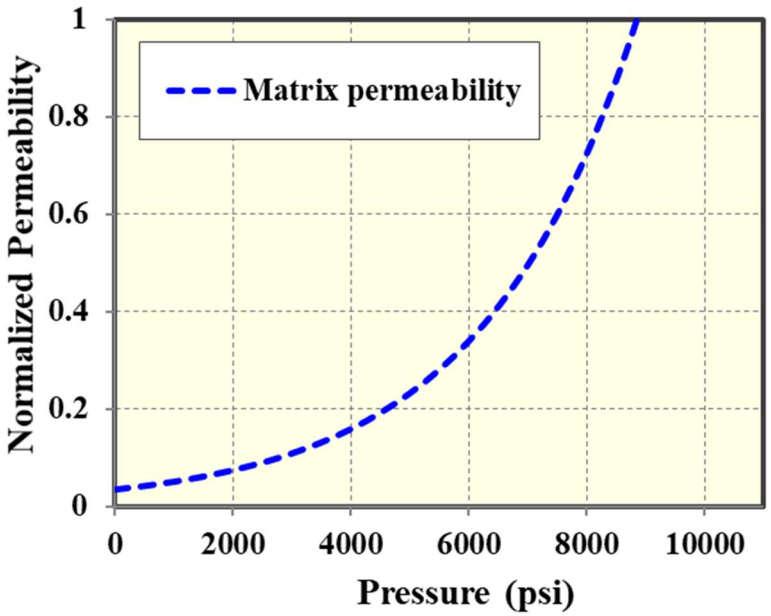


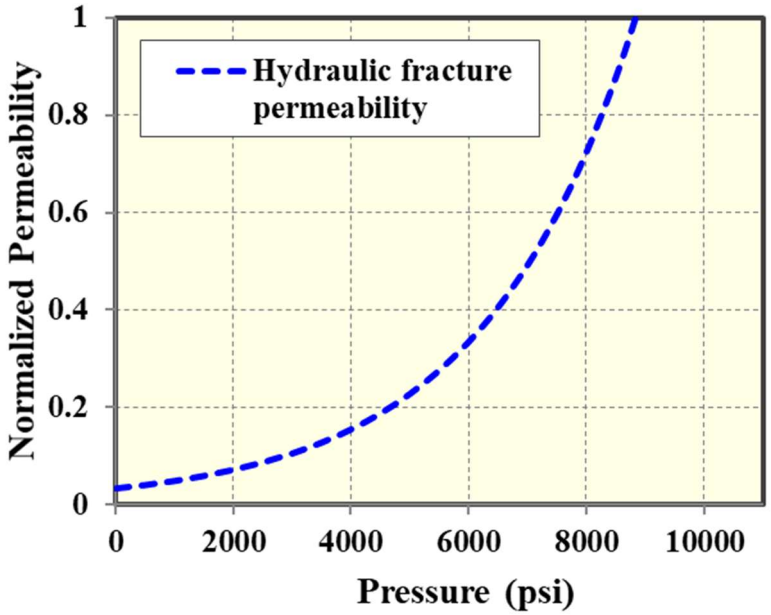
Figure 6.1: Reservoir model including 58 hydraulic fractures (in blue) modeled by EDFM method in a horizontal well (in red).

Langmuir isotherm parameters	**Value**	**Unit**
Inverse-pressure parameter for the Langmuir isotherm model, CH_4	0.000534	1/psi
Maximum moles of adsorbed component per unit mass of rock, CH_4	0.060148	gmole of component/lb of rock
Rock density	162.318	lb/ft^3

Table 6.3: Langmuir isotherm parameters for modeling the gas desorption effect



(a) Pressure-dependent matrix permeability



(b) Pressure-dependent fracture permeability

Figure 6.2: Pressure-dependent normalized permeability for (a) Matrix permeability and (b) Fracture permeability to capture the geomechanical effect of the permeability reduction when reservoir is depleted.

6.3 HISTORY MATCHING

In this section, we performed history matching for two-phase flow of gas and water in shale gas reservoirs. In addition, we investigated the effect of natural fracture existence by performing two cases including hydraulic fractures only and hydraulic fractures with natural fractures. We had a historical production of 724 days including gas flow rate, water flow rate, and wellhead pressure. For history matching, we used gas flow rate, water flow rates and bottomhole pressure (BHP) calculated from wellhead pressure (WHP) due to wellbore pressure loss. We assumed the constant pressure loss from wellhead to bottomhole.

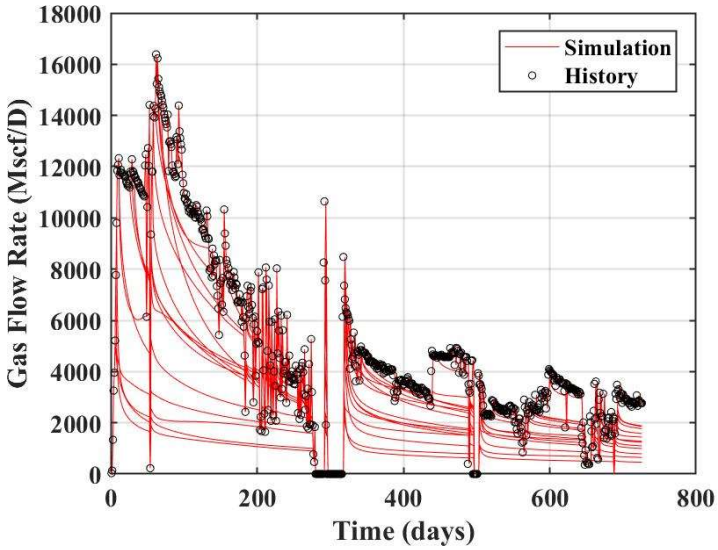
6.3.1 Case 1 Hydraulic Fractures Only (No NF)

For the case with only hydraulic fractures, we identified the uncertain parameters and their ranges as summarized in Table 6.4. We can categorize the uncertain parameters into two groups including first, reservoir properties including matrix permeability, matrix water saturation and relative permeability curves parameters, and second, hydraulic fractures properties including fracture height, fracture half-length, fracture conductivity, fracture water saturation and fracture width. We determined the initial ranges of each parameter based on our initial knowledge or prior knowledge about this well and reservoir. Based on our experience, the success of history matching significantly depends on the reasonable parameter values in order to mimic the production data that we are matching. We cannot achieve the reasonable matching quality with the wrong parameter range inputs, even though, an accurate and efficient history-matching algorithm is already used. In case, the good history matching quality cannot be achieved, the first thing to review are parameter inputs and then adjust them.

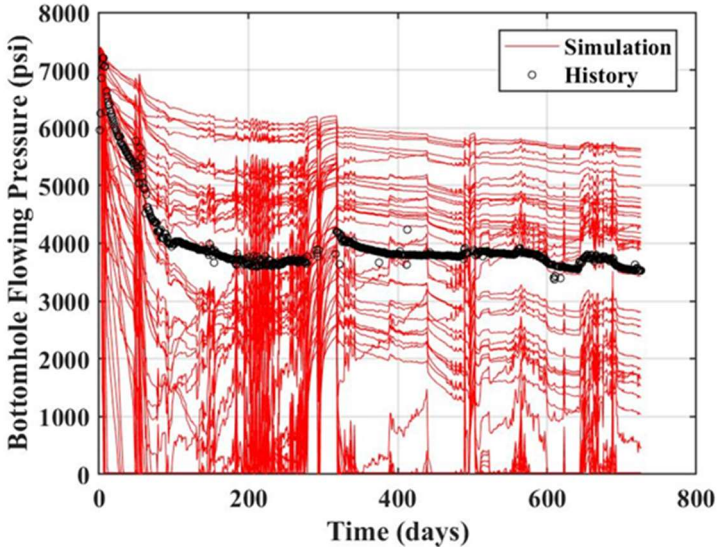
Then, we used Latin Hypercube design to populate 50 initial cases scattering the entire parameter domain. This will benefit the proxy model construction to achieve the better quality rather than using random sampling as Monte Carlo method. The simulation results of 50 initial cases from Latin Hypercube design are shown in Figure 6.3. We could observe that the simulation results cover the historical data and their shapes are similar to production history. This means we can possibly find history matching solutions in our initial uncertain parameter ranges. On the other hand, if the simulation results do not cover the historical data or the shapes of production profile are totally different, we need to review our initial uncertain parameter range inputs and revise them.

Uncertain Parameter	**Unit**	**Min Value**	**Max Value**
Matrix Permeability (km)	md	0.0001	0.04
Fracture Height	ft	16	115
Fracture Half-length (Xf)	ft	200	1000
Fracture Conductivity	md-ft	1	100
Fracture Water Saturation	-	0.50	0.90
Matrix Water Saturation	-	0.30	0.50
Exponent of krg	-	1	4
Endpoint of krw	-	0.5	1
Exponent of krw	-	1	4
Fracture Width	ft	0.1	1

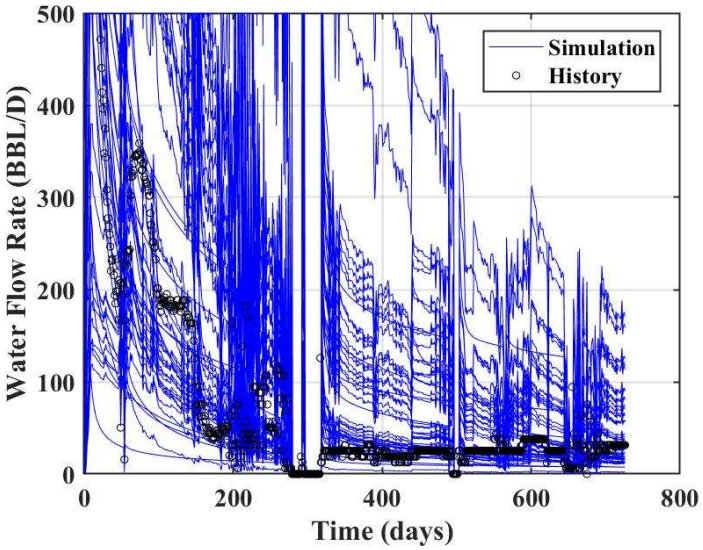
Table 6.4: Summary of the uncertain parameters of reservoir and fractures properties



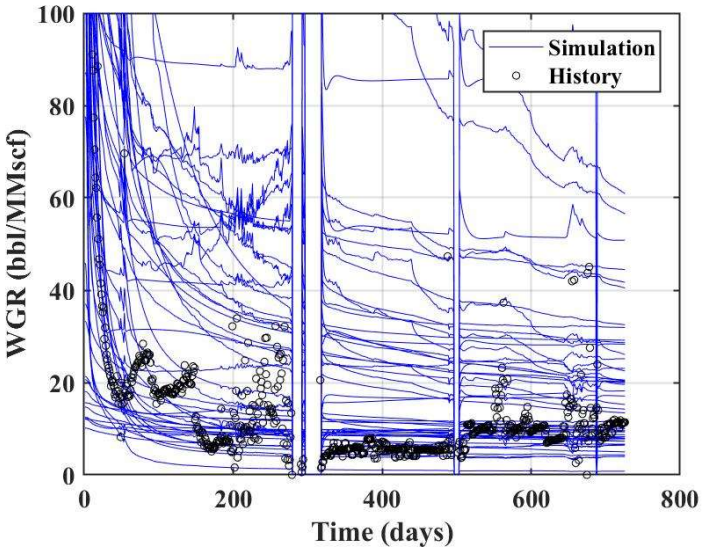
(a) Gas flow rate as simulation constraint



(b) Flowing BHP



(c) Water flow rate



(d) Water-gas ratio

Figure 6.3: Comparison between simulation results and historical data of the cases from Latin Hypercube design (Iteration 1): (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate, and (d) Water-gas ratio (WGR).

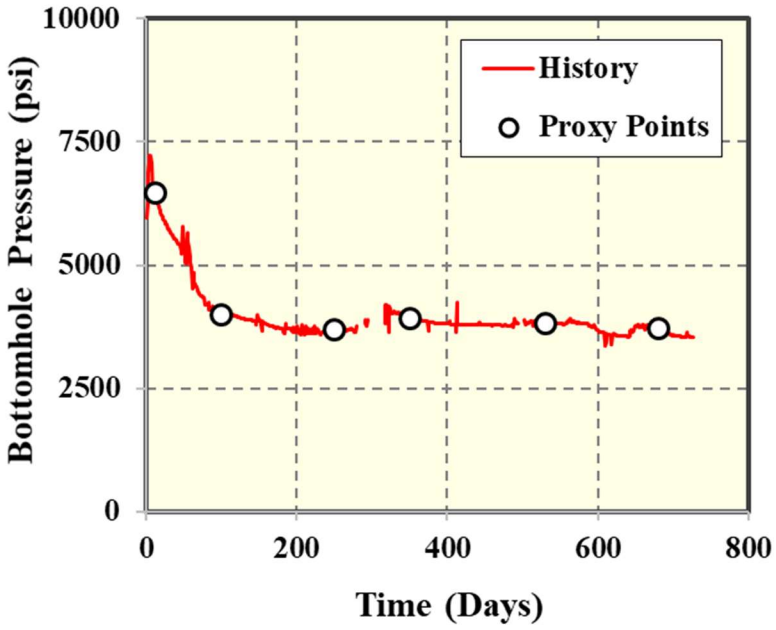
Instead of constructing one proxy model for each response parameter, we used multiple proxy models to mitigate the nonlinearity of the problem. Therefore, we had to define the proxy points to construct proxy models (NNs in this case study for proxy-based MCMC algorithm). The global objection function will be calculated from many points that we selected based on assigned weight values. In an ideal case, we could choose many proxy points as long as the computational effort allows. However, adding unnecessary proxy points would hinder the performance of history matching algorithm. This means we need to achieve the acceptable accuracy with the optimized computational effort. The criterion is to choose proxy points that capture the major trend of production data. In this case study, we used total 12 proxies from different dates in production data. This included 6 proxies for BHP at day 12, 100, 250, 350, 530 and 680 (Figure 6.4a) and 6 proxies for water gas ratio (WGR) at 12, 50, 120, 180, 400 and 600 (Figure 6.4b). All 12 proxies captured the major trend of BHP and WGR. The reason we used WGR instead of water flow rate to represent water phase is that water-gas ratio behavior in shale gas reservoirs is predictable. The trend will start from the high water-gas ratio due to initial flow back period from fractures then drop to some constant WGR reflecting matrix response. This requires a smaller number of proxy points to capture the WGR trend due to the constant part after flow back period, as illustrate in Figure 6.4b. Moreover, the water-gas ratio can directly be related to relative permeability curves and water saturation for multiphase flow behavior.

We used NN as a proxy model to predict the values without running actual simulations during performing NN-MCMC algorithm. The NN-MCMC algorithm progressed iteratively. In each iteration, the algorithm proposed new 25 cases as history matching solution candidates to be validated with the actual reservoir simulator. Then, new simulation results were included into the training data set and then re-train the NNs. The

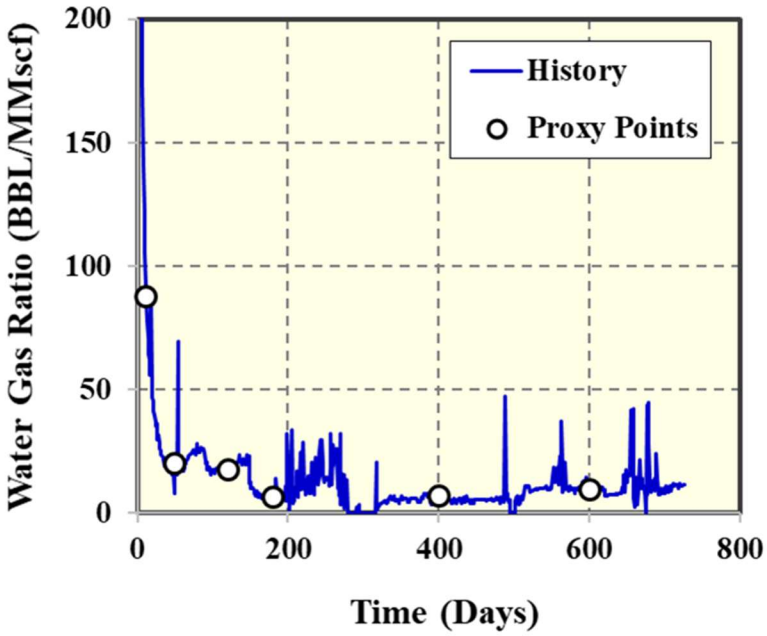
stopping criteria of the workflow are either the proxy convergence or maximum number of iterations. In this study, we used proxy convergence value of 5% difference between current proxy and previous proxy. The workflow stopped at the maximum of 12 iterations in this case.

After the AHM workflow stopped and all simulation runs were completed, we screened history-matching solutions from all runs of all iterations. We defined the history matching criteria of 10 % and 40 % for BHP and WGR responses based on our decision for the least acceptable matching quality. The history matching solutions must pass these two criteria, or the intersection of two response parameters as shown in Figure 6.5. The criteria for WGR is higher due to more acceptable range of water phase and also their low values naturally create a higher value of relative error percentage. For example, if the historical WGR is 5 BBL/MMscf and the simulation WGR is 10 BBL/MMscf. Because we used relative error to calculate WGR error objective function, as shown in Equation 3.18, the discrepancy between historical data and simulation result would be 100%. Even though the absolute value difference is only 5 BBL/MMscf.

Figure 6.6 illustrates the number of history matching solutions vs. iteration number. We obtained 89 solutions from the total of 325 simulation runs, which is around 27%. One thing to highlight is that more solutions were found at the later iteration because the quality of NN is better - more actual data points from more simulation runs are included in the NN training. All simulation results of 89 history matching solutions are shown in Figure 6.7, and the simulation results of the best match case for hydraulic fractures only case (No NF) are illustrated in Figure 6.8, including BHP, gas flow rate, water flow rates and WGR.



(a) 6 proxies of BHP



(b) 6 proxies of WGR

Figure 6.4: Multiple proxies of 12 including (a) 6 proxies of BHP and (b) 6 proxies of WGR.

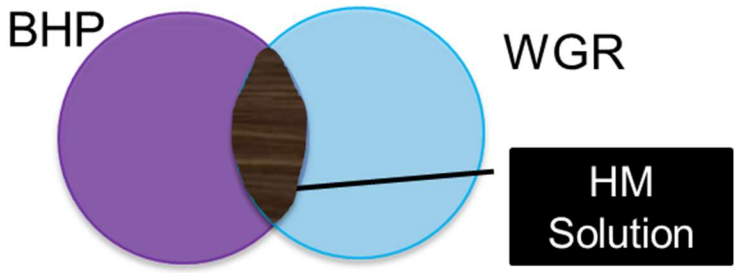


Figure 6.5: History matching solutions are screen by the criteria of 10 and 40 for each objective function of BHP and WGR, respectively.

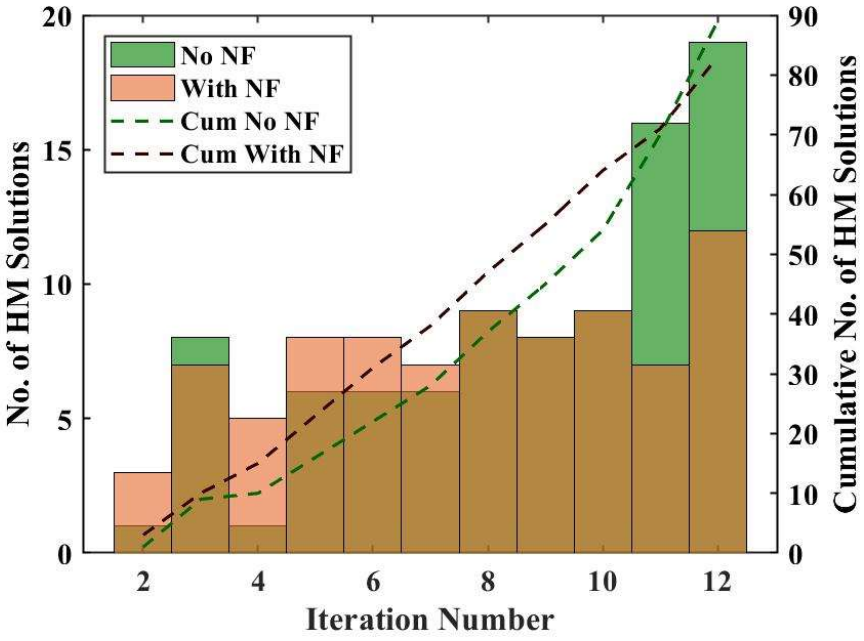
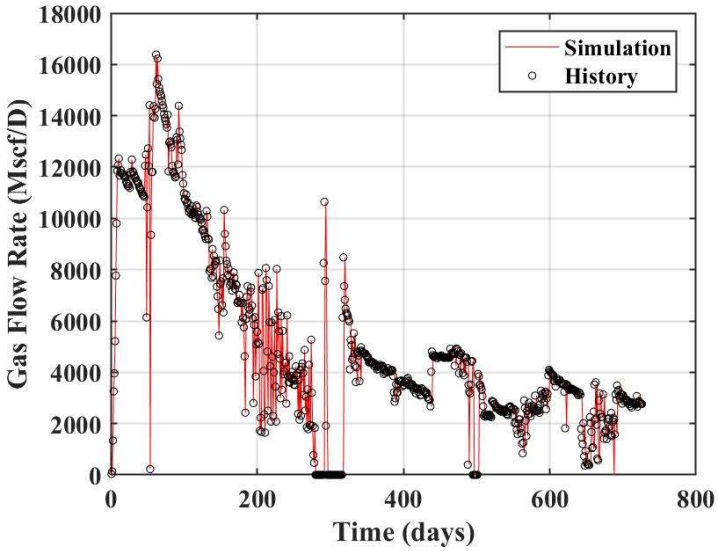
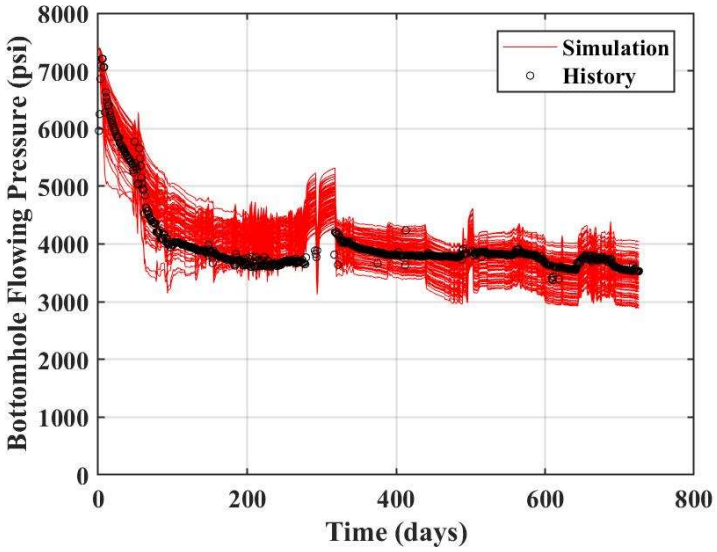


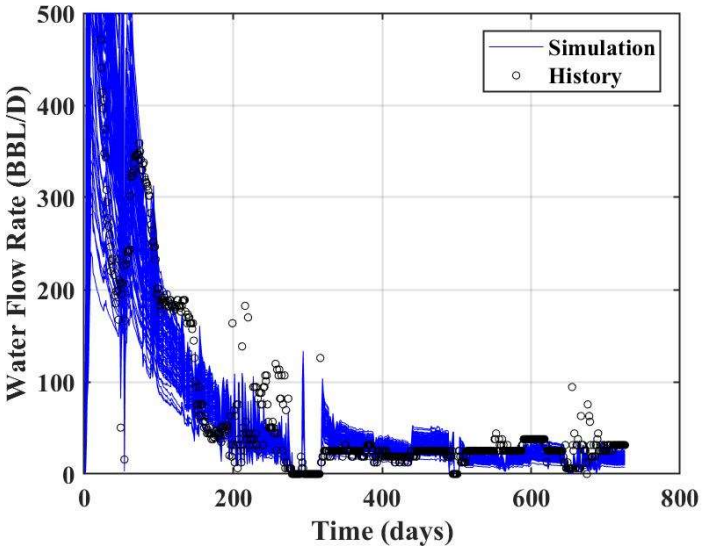
Figure 6.6: Number of history matching solutions vs. iteration number for case 1 hydraulic fractures only (No NF) and case 2 hydraulic fractures and natural fractures (With NF).



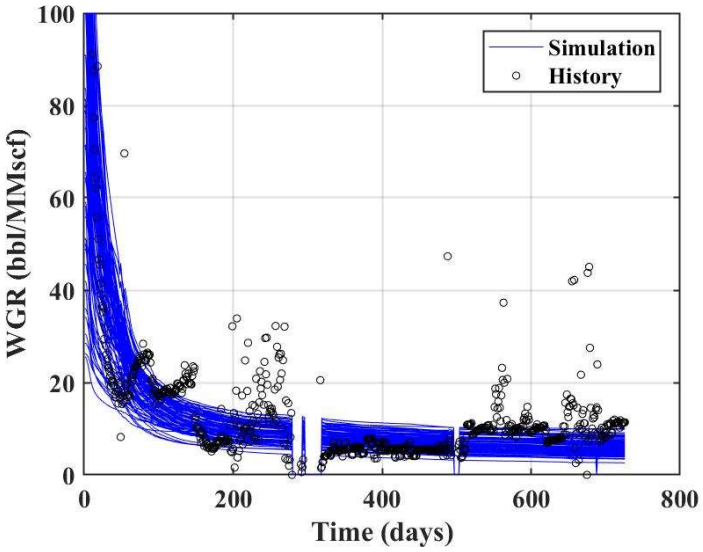
(a) Gas flow rate as simulation constraint



(b) Flowing BHP

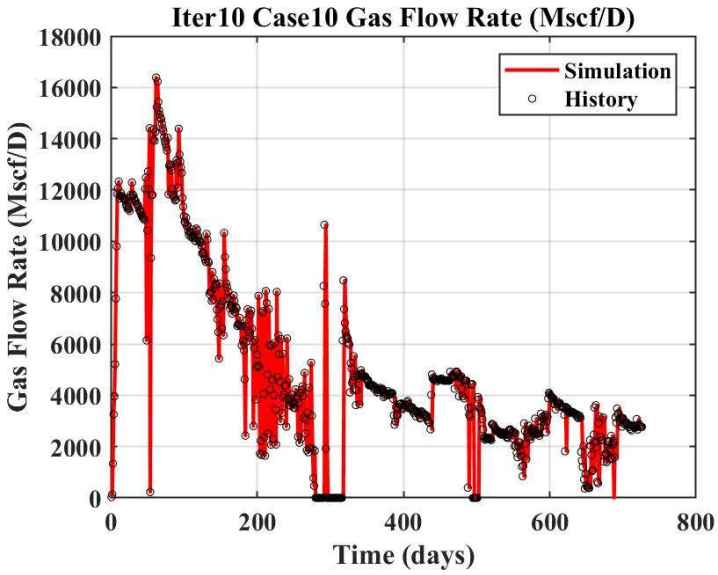


(c) Water flow rate

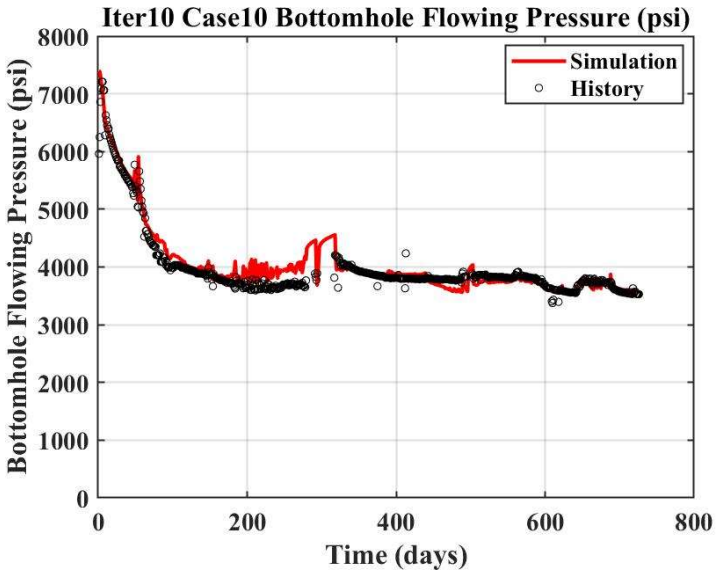


(d) Water-gas ratio

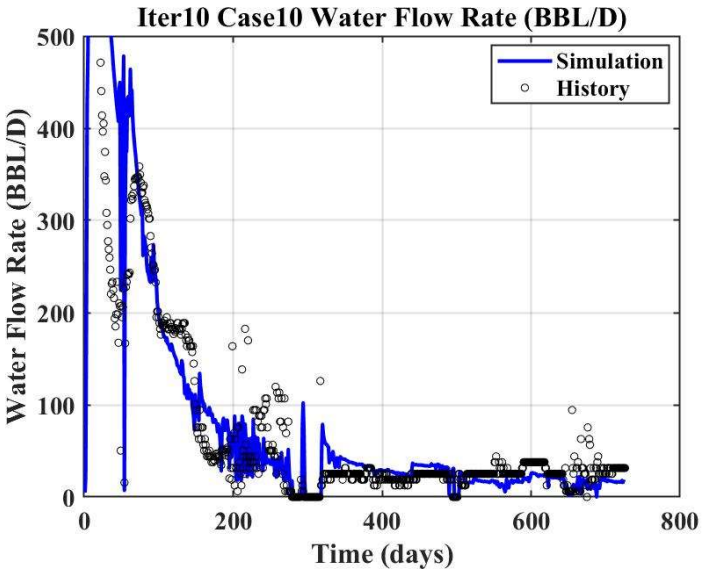
Figure 6.7: Comparison between simulation results of 89 history matching solutions and historical data of case 1 with hydraulic fractures only (No NF): (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) Water-gas ratio (WGR).



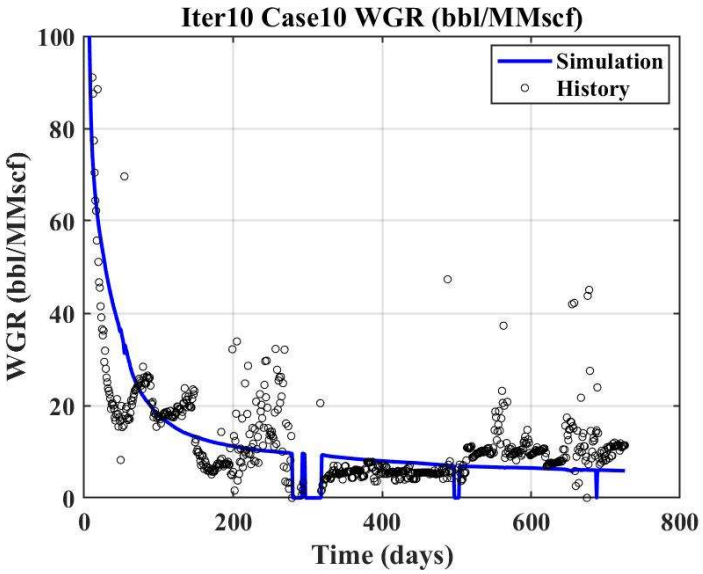
(a) Gas flow rate as simulation constraint



(b) Flowing BHP



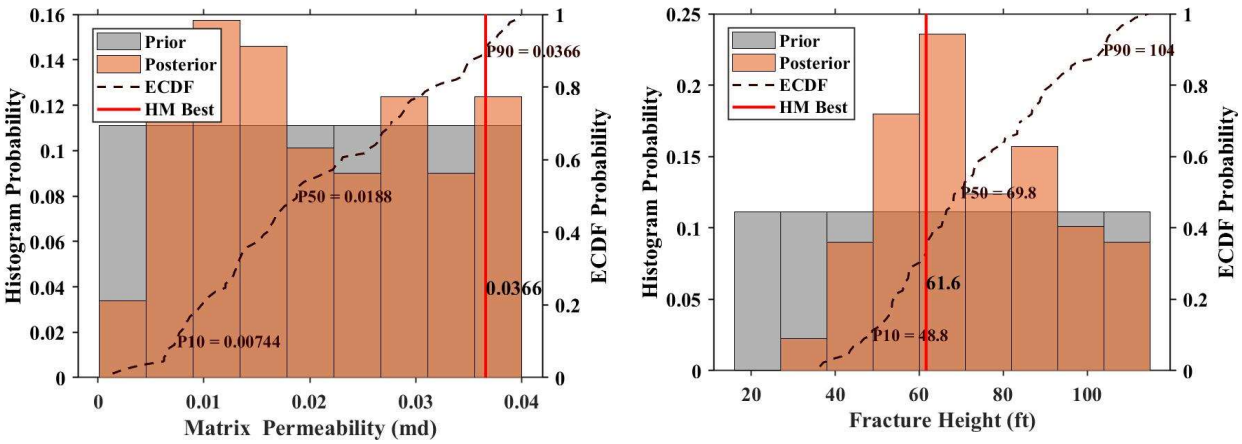
(c) Water flow rate



(d) Water-gas ratio

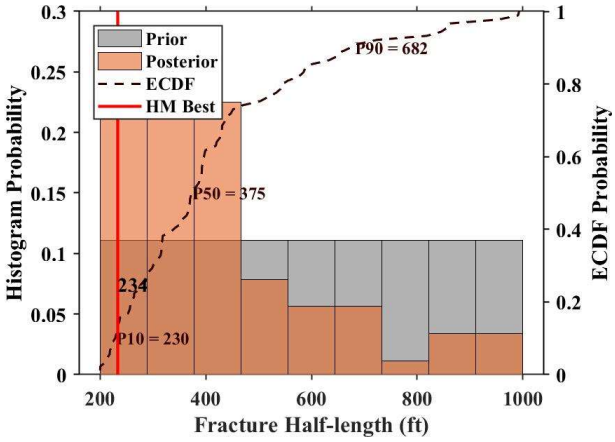
Figure 6.8: Comparison between simulation results of the best match case (Iteration 10 Case 10) and historical data of the case with hydraulic fractures only (No NF): (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) Water-gas ratio (WGR).

One thing we can obtain from history matching workflow is the posterior distribution of uncertain parameters from 89 history-matching solutions. As shown in Figure 6.9, the prior uniform distribution we specified at the beginning has changed to the posterior distribution. This means we can characterize reservoir and fractures properties from historical production data in a probabilistic manner. Not just one realization of history matching solution is obtained as the traditional method and we still are uncertain what this case would locate in the uncertainty range. While the initial range of matrix permeability 0.0001-0.04 md, exponent of krg of 1-4 and endpoint of krw 0.5-1 can make combinations with other uncertain parameters to produce history matching solutions, we can characterize other uncertain parameters; for example, fracture height mostly in the range of 40-100 ft, fracture half-length in the lower bond and fracture conductivity in the higher side near 50-80 md-ft tend to produce as history matching solutions. In addition, fracture water saturation in the range of 0.55-0.8, matrix water saturation of 0.35-0.45, the exponent of krw in the higher end of 4 and fracture width below 0.8 ft are likely to match the values of historical WGR trend.

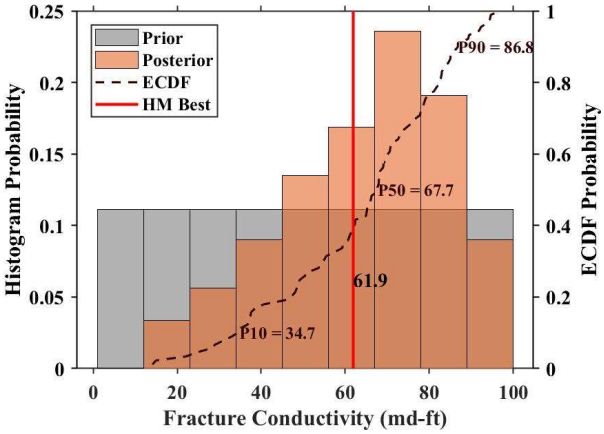


(a) Matrix permeability

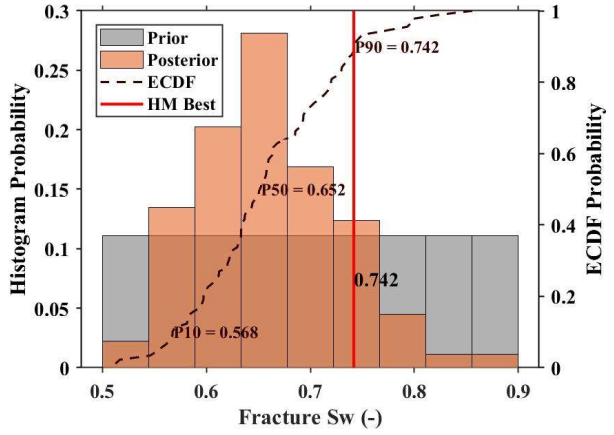
(b) Fracture height



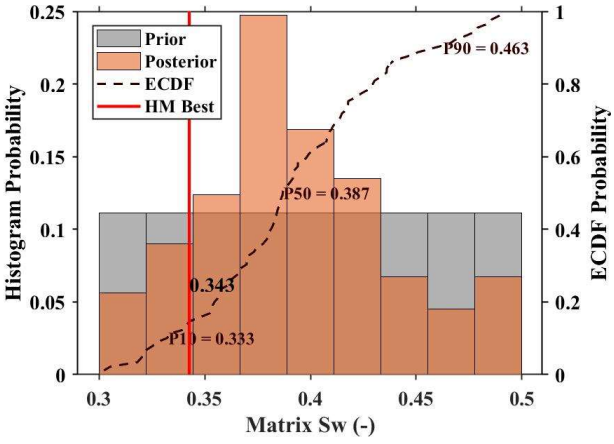
(c) Fracture half-length



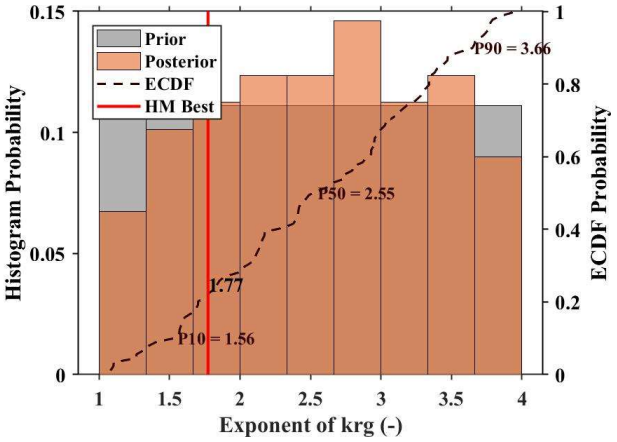
(d) Fracture conductivity



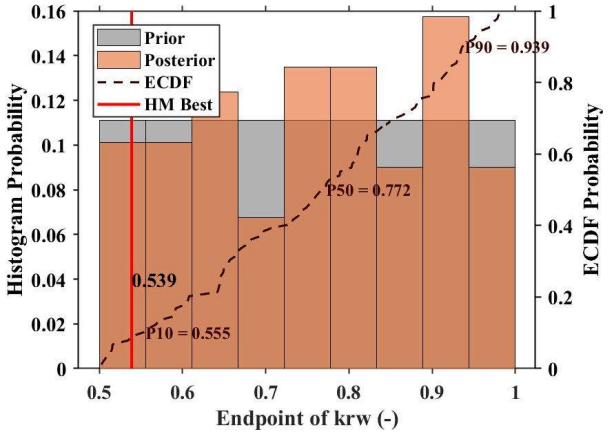
(e) Fracture water saturation



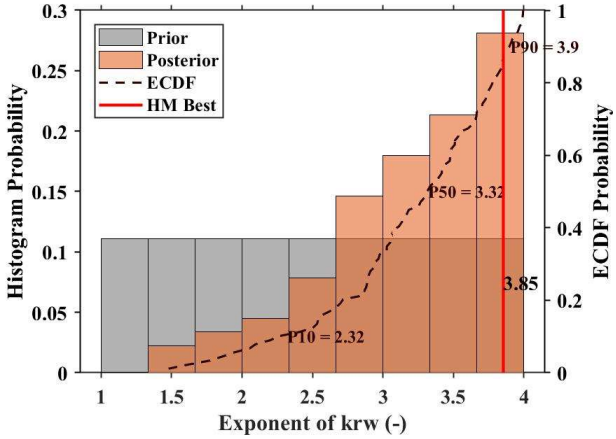
(f) Matrix water saturation



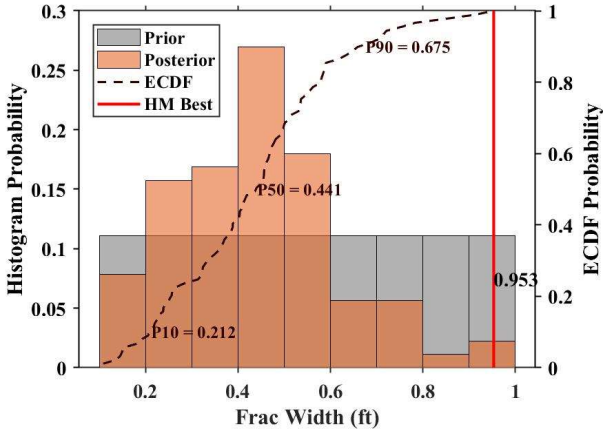
(g) Exponent of k_{rg}



(h) Endpoint of k_{rw}



(i) Exponent of k_{rw}



(j) Fracture width

Figure 6.9: Prior and posterior distribution of case 1 with hydraulic fractures only (No NF) with P10-50-90 values and best match case.

6.3.2 Case 2 Hydraulic Fractures and Natural Fractures (With NF)

The second case we investigated is the case of hydraulic fractures and natural fractures. The objective is to evaluate the effect of natural fractures existence on the history matching solutions and production forecast. All reservoir model information and uncertain parameter ranges are same as in case 1 (Table 6.1-6.3). In this case, we added uncertain parameters of natural fractures properties including number of natural fractures, natural fractures length and natural fractures conductivity as shown in Table 6.5. The number and the length of natural fractures we used are from our prior knowledge of the natural fractures length per area in this reservoir. The maximum of natural fracture conductivity is assumed ten times lower than the maximum of hydraulic fractures conductivity. These three uncertain parameters normally have more impact on the flow behavior and overshadow other natural fractures parameters. Therefore, we used deterministic values of all other natural fractures parameters such as theta, dip angle and natural fracture width, as summarized in Table 6.6. We specified two sets of natural fractures with different theta. To illustrate, Figure 6.10 shows hydraulic fractures with natural fractures modeled by EDFM. With the effect from natural fractures, our hypothesis is that the natural fractures intersection with themselves and hydraulic fractures could enhance more connection to the production well. This means the existence of natural fractures could affect the history matching solutions of other fractures and reservoir uncertain parameters and production forecast, but we still do not know how much this impact would be until the history matching is performed.

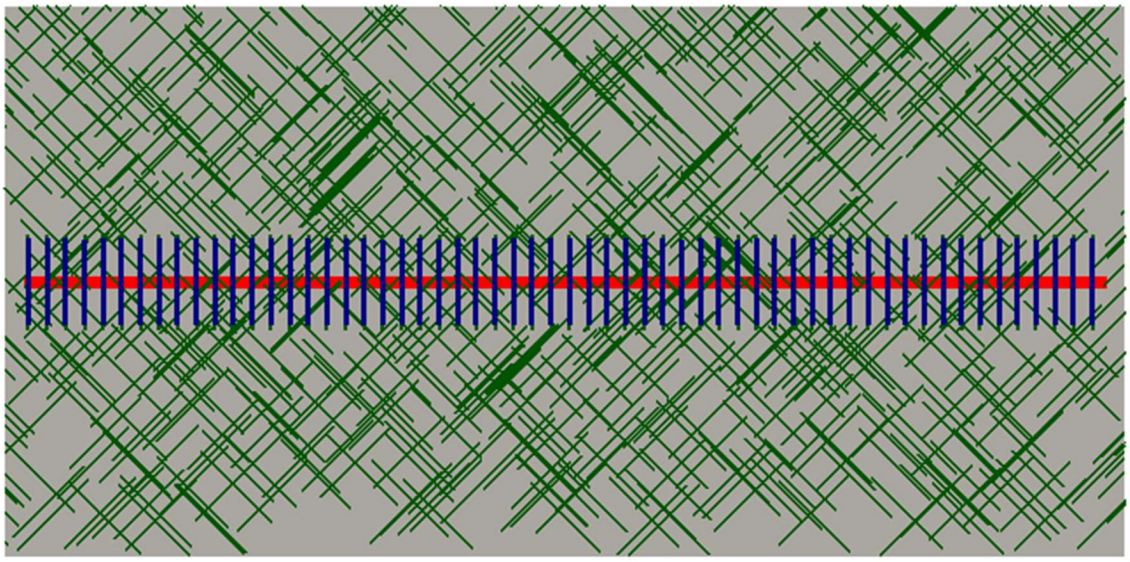


Figure 6.10: The horizontal well (in red) with hydraulic fractures of 58 clusters (in blue) and two sets of natural fractures modeled by EDFM.

Uncertain Parameter	**Unit**	**Min Value**	**Max Value**
Total number of NF	-	200	1200
NF Length (same for both NF sets)	ft	200	800
NF Conductivity	md-ft	1	10

Table 6.5: Summary of the uncertain parameters of natural fractures properties

Fixed Parameter	Unit	Value
Number of NF Set	-	2
NF Height	ft	As HF height
NF Theta	degree	NF set 1: 45 NF set 2:135
NF Dip Angle	degree	90

Table 6.6: Summary of the fixed parameters of natural fractures properties

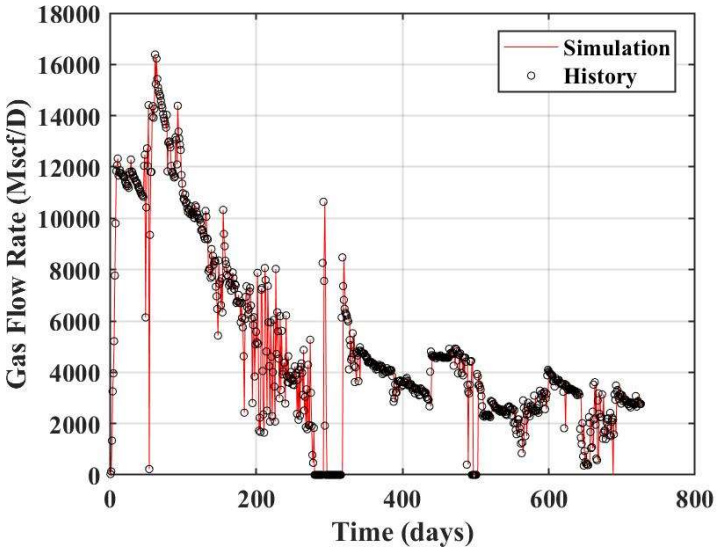
The history matching process progressed the same as case 1. In this case, the history matching workflow also stopped at the maximum iteration of 12. With the same history matching solution screening criteria, the workflow found 83 history matching solutions from total 325 simulation runs (around 26%) compared to 89 history matching solutions of case 1 with only hydraulic fractures (Figure 6.6). The lower number of history matching solutions is obviously from the higher degree of freedom including natural fractures uncertain parameters.

All simulation results of 83 history-matching solutions are shown in Figure 6.11, including gas flow rate, water flow rates and WGR. Figure 6.12 also illustrates the simulation results of best match case of the case with natural fractures. Similarly, we can obtain posterior distribution of uncertain parameters, especially natural fractures parameters, together with best match case values and P10-50-90 values, as shown in Figure 6.13. The comparison between prior and posterior distribution of uncertain parameters for both case 1 (No NF) and case 2 (With NF) are presented in Figure 6.14.

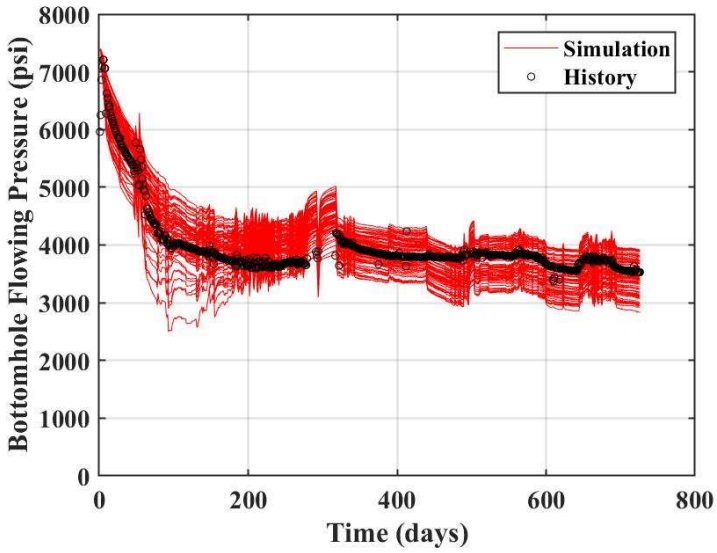
When natural fractures exist together with hydraulic fractures, the posterior distribution of some uncertain parameters significantly changed from the case with only

hydraulic fractures. While there is no significant change for matrix permeability and fracture half-length, we observed the significantly lower fracture height, lower fracture conductivity, higher fracture water saturation because more fluid flow is enhanced by natural fractures. For relative permeability, we see the lower exponent of krg and lower exponent of krw or less curvature on both krg and krw. These changes are possibly to compensate lower flow from other uncertain parameters. Lastly, the uncertainty of natural fractures can be reduced from the prior distribution by using the AHM workflow. As shown in Figure 6.13 k, l and m, we can observe that natural fractures conductivity and natural fracture length did not change significantly from prior distribution while we see the significant change in the posterior distribution of number of natural fractures from the initial prior distribution. The higher end values tend to provide more history matching solutions. This implies that, for this case, number of natural fractures has more impact on the production, than natural fractures length and conductivity because any value of natural fractures length and conductivity can yield history matching solutions

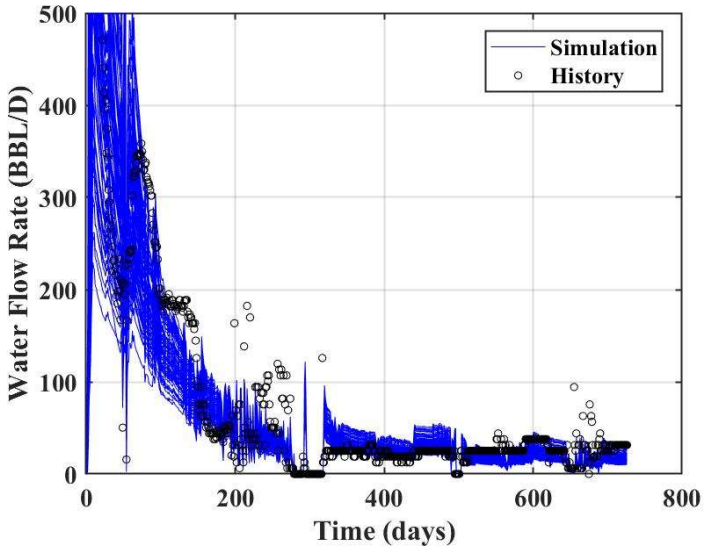
The summary of uncertain parameters values of the best match cases for both case 1 with hydraulic fractures only (No NF) and case 2 with hydraulic fractures and natural fractures (With NF) are given in Table 6.7. The pressure distributions for both cases are also depicted in Figure 6.15. We can see the difference in pressure profile of uniform and non-uniform depleted area in x-y plane for case 1 and case 2, respectively. One thing to note is that these depleted areas in x-y between two cases are not obviously different because the fracture half-length posterior distribution of two cases (Figure 6.14c) are not significantly different based on history matching. In fact, the main difference should be from fracture height (Figure 6.14b) or in z direction, which is previous discussed.



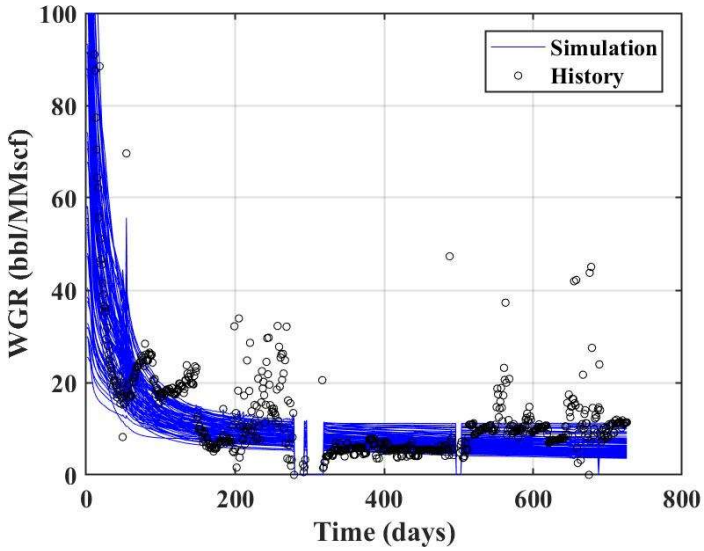
(a) Gas flow rate as simulation constraint



(b) Flowing BHP

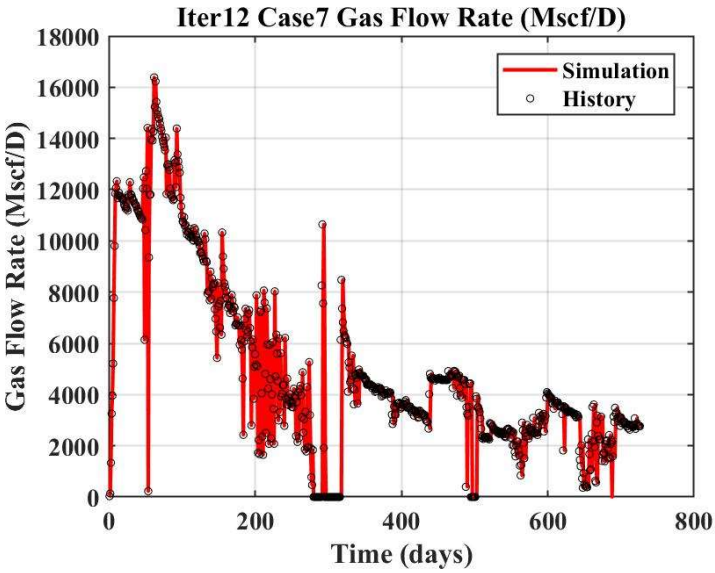


(c) Water flow rate

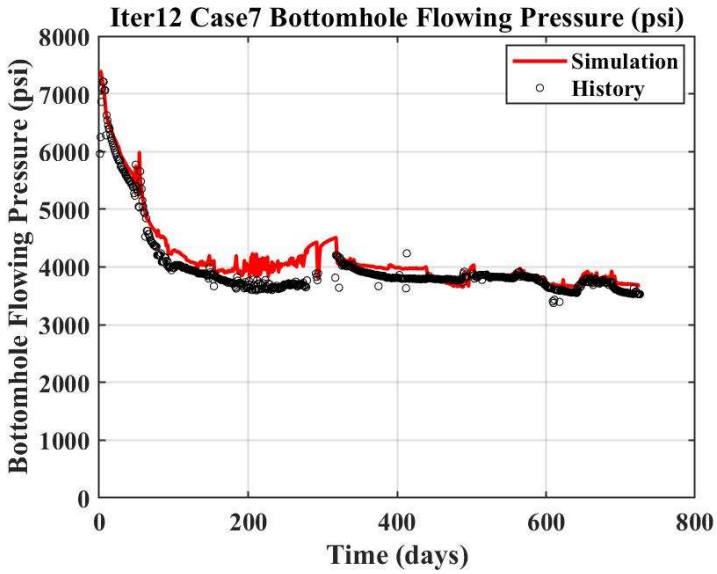


(d) Water-gas ratio

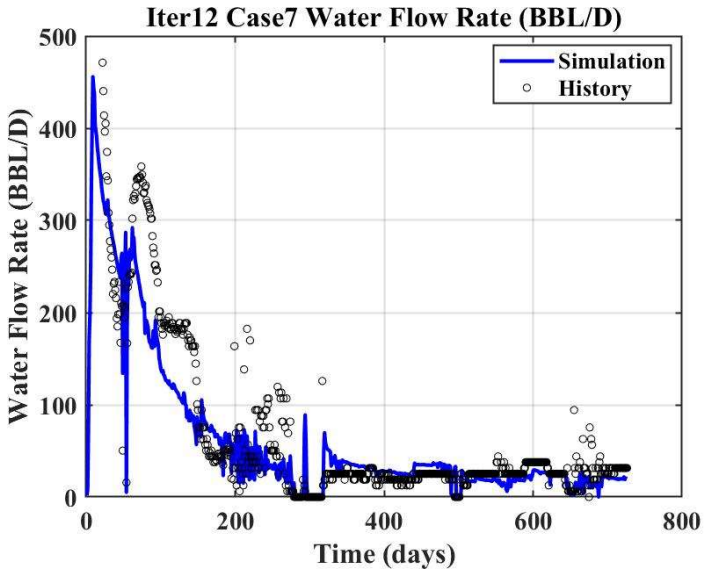
Figure 6.11: Comparison between simulation results of 83 history matching solutions and historical data of case 2 with hydraulic fractures and natural fractures (With NF): (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) Water-gas ratio (WGR).



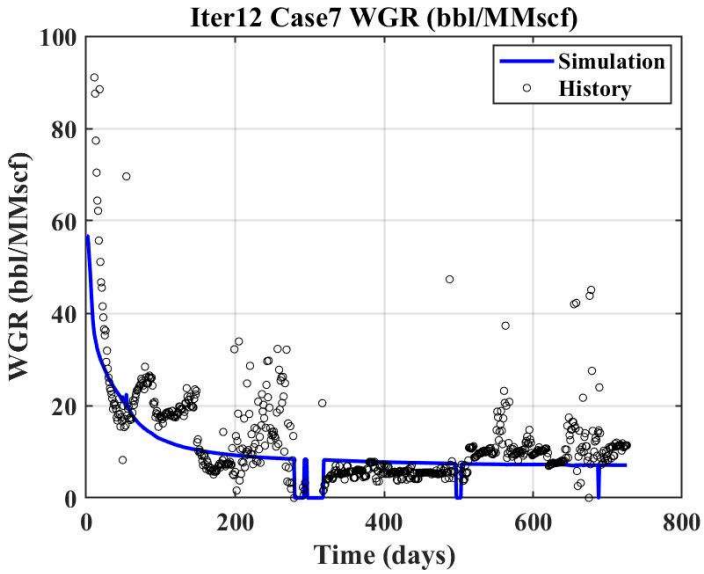
(a) Gas flow rate as simulation constraint



(b) Flowing BHP

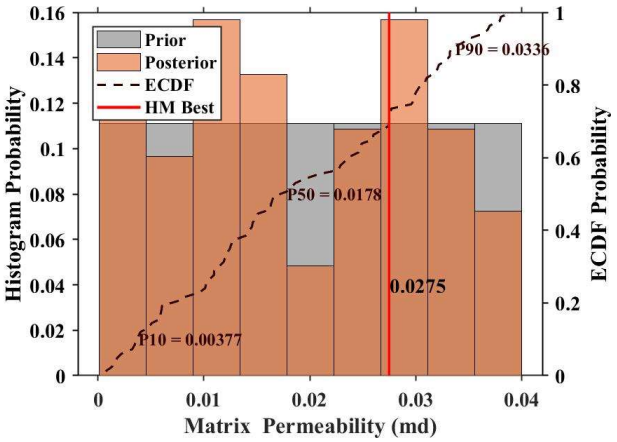


(c) Water flow rate

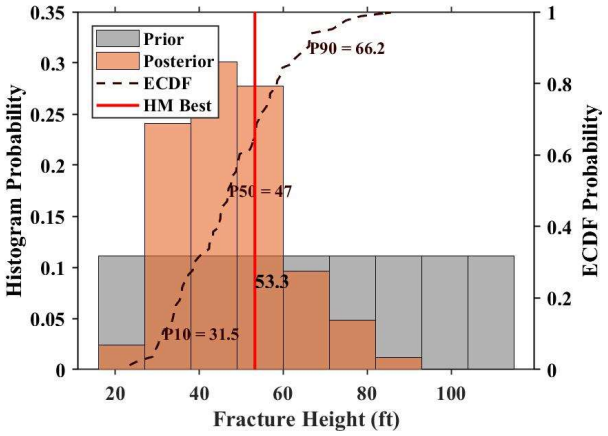


(d) Water-gas ratio

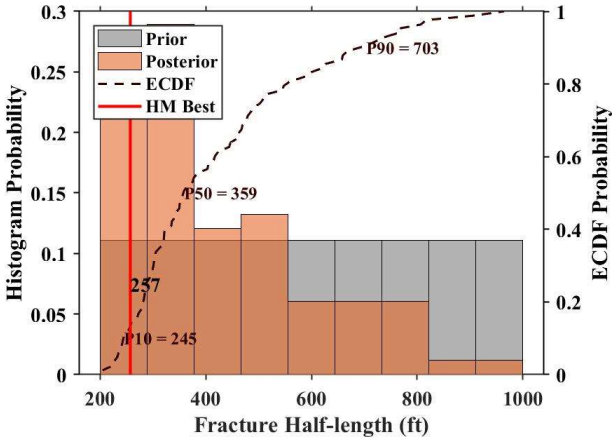
Figure 6.12: Comparison between simulation results of the best match case (Iteration 12 case 7) and historical data of case 2 with hydraulic fractures and natural fractures (With NF): (a) Gas flow rate, (b) Flowing BHP, (c) Water flow rate and (d) Water-gas ratio (WGR).



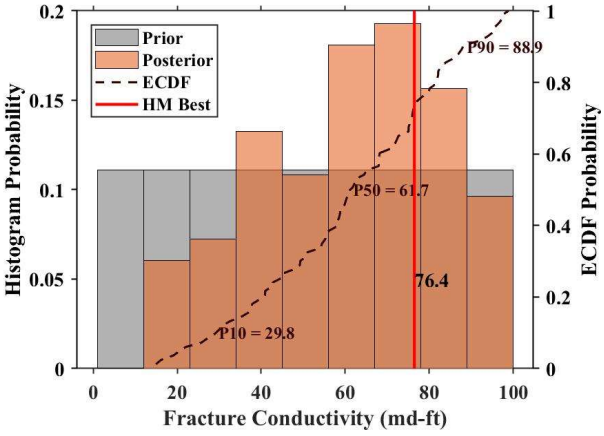
(a) Matrix permeability



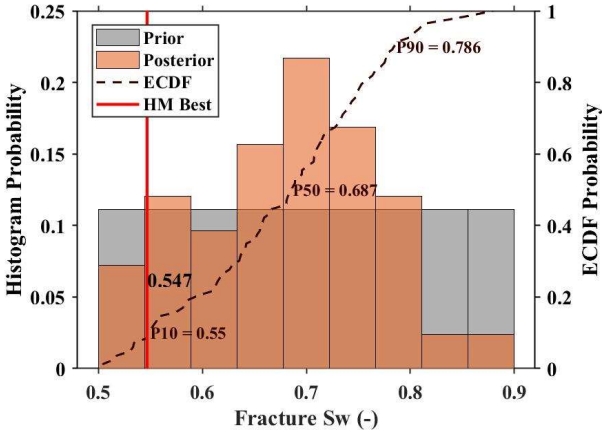
(b) Fracture height



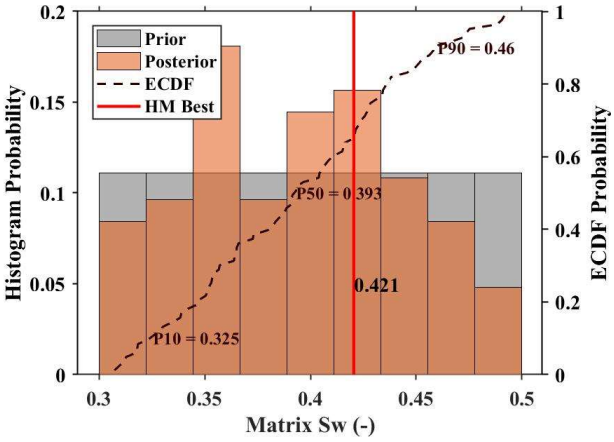
(c) Fracture half-length



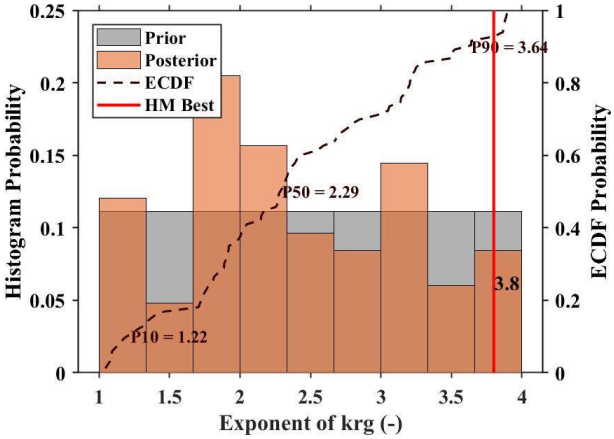
(d) Fracture conductivity



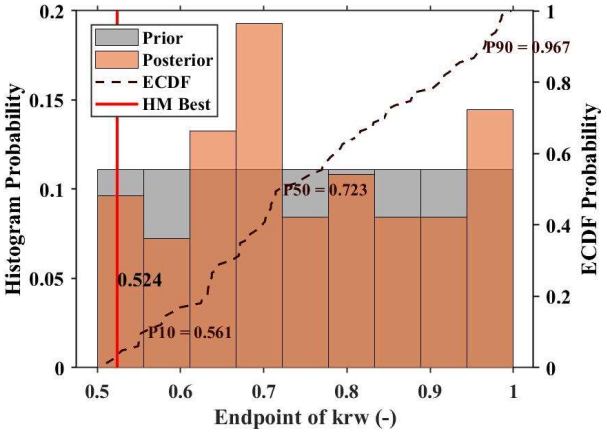
(e) Fracture water saturation



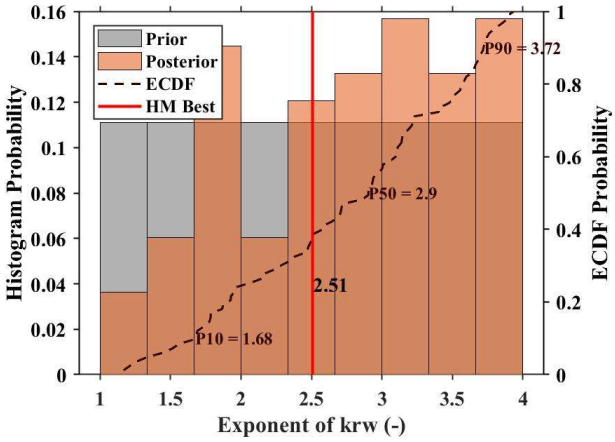
(f) Matrix water saturation



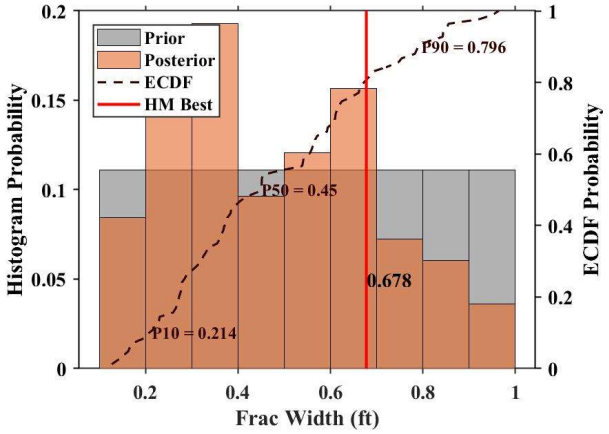
(g) Exponent of k_{rg}



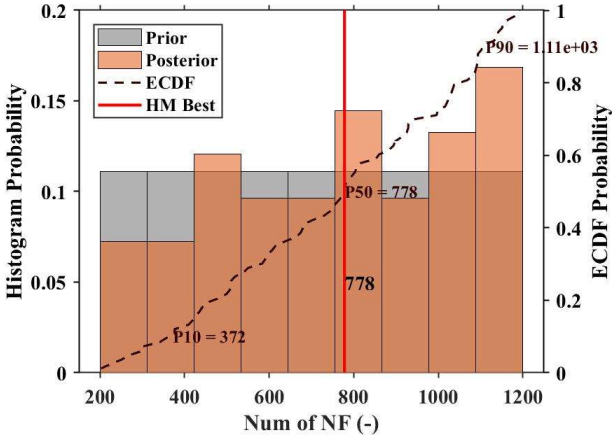
(h) Endpoint of k_{rw}



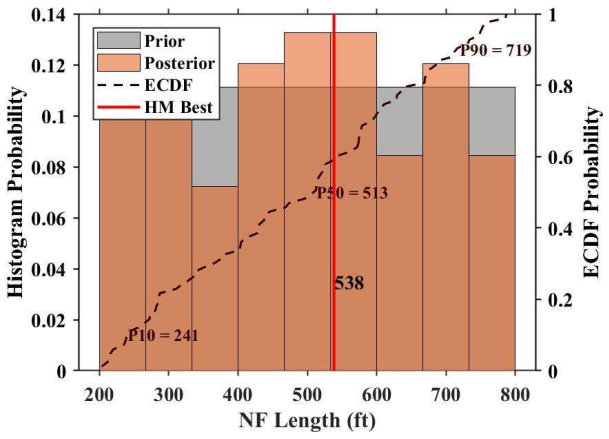
(i) Exponent of k_{rw}



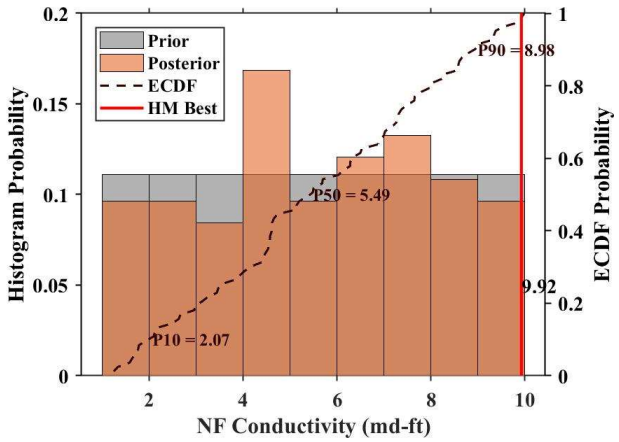
(j) Fracture width



(k) Number of NF

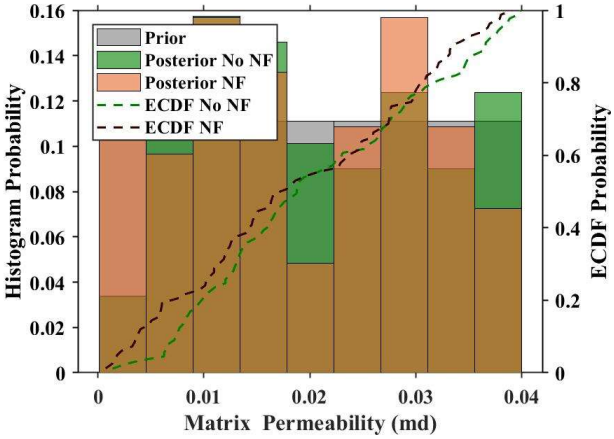


(l) NF length

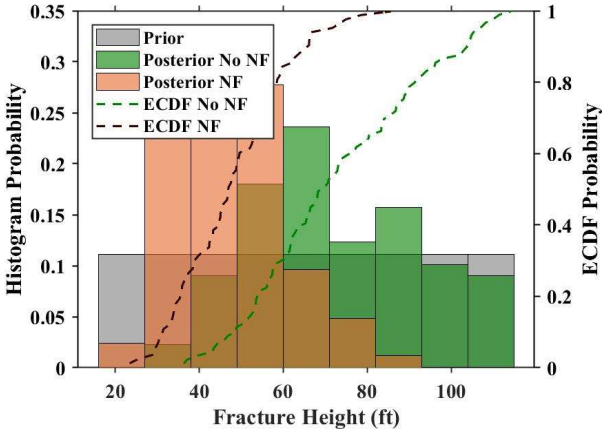


(m) NF conductivity

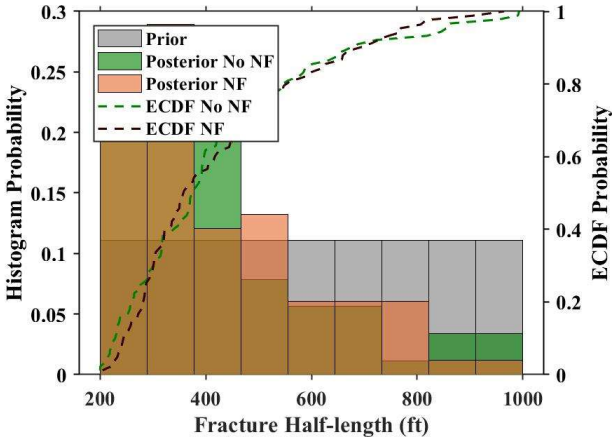
Figure 6.13: Prior and posterior distribution of case 2 with hydraulic fractures and natural fractures (With NF) with P10-50-90 values and best match case.



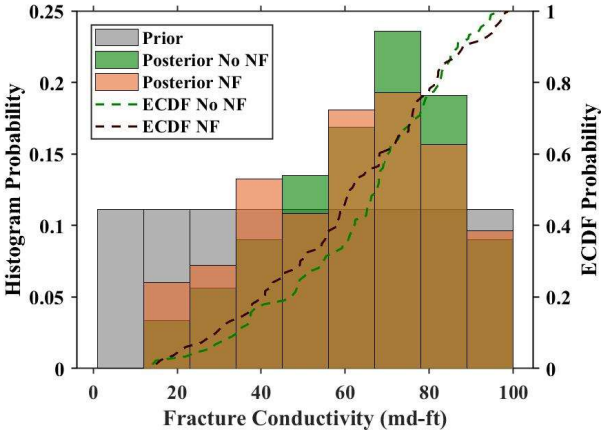
(a) Matrix permeability



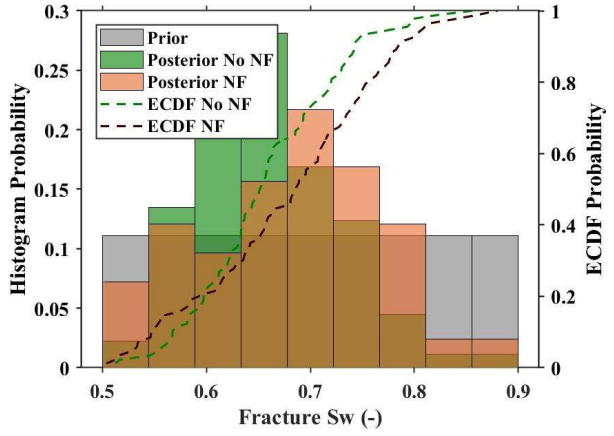
(b) Fracture height



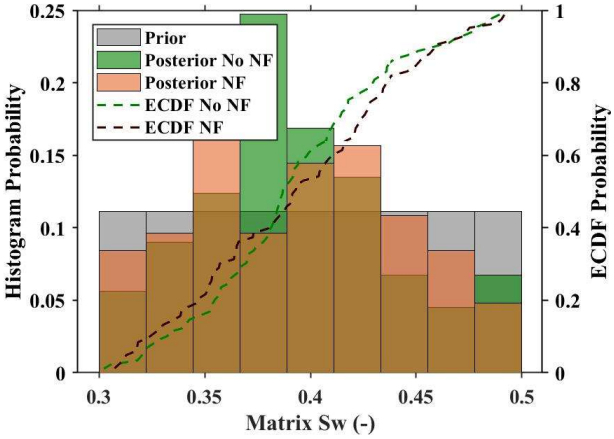
(c) Fracture half-length



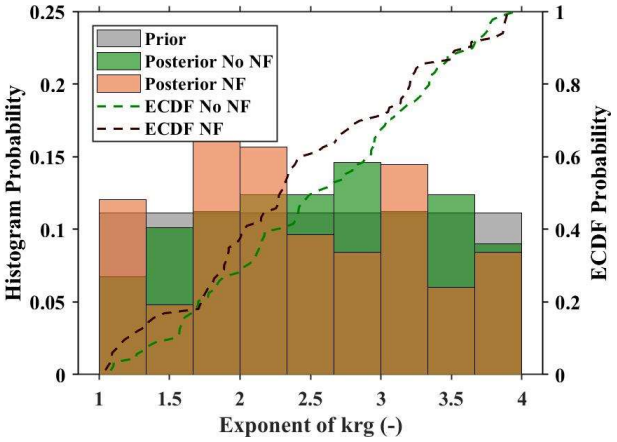
(d) Fracture conductivity



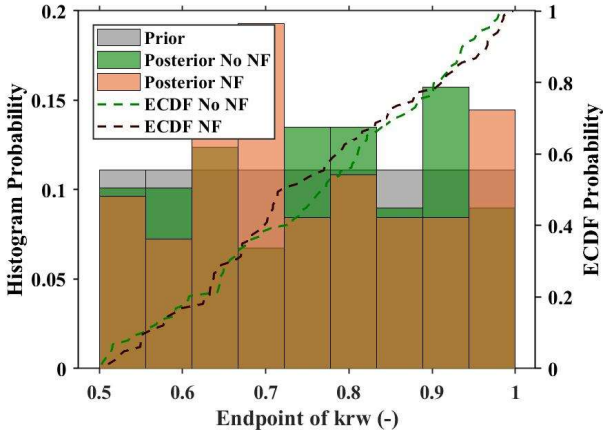
(e) Fracture water saturation



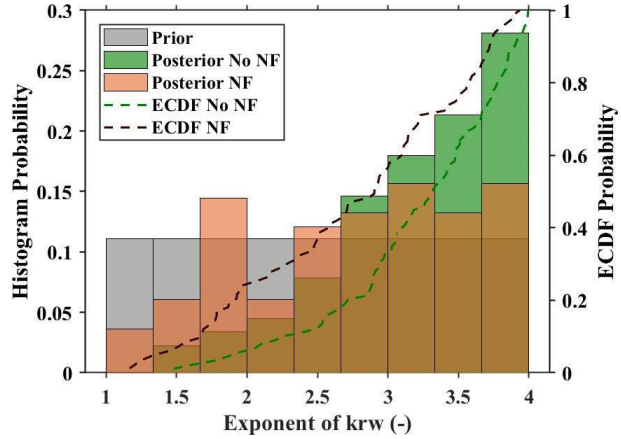
(f) Matrix water saturation



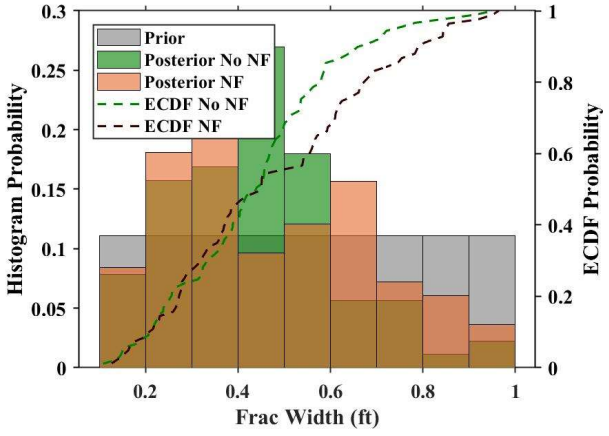
(g) Exponent of k_{rg}



(h) Endpoint of k_{rw}



(i) Exponent of k_{rw}



(j) Fracture width

Figure 6.14: Prior and comparison of posterior distribution between case 1 with hydraulic fractures only (No NF) and case 2 with hydraulic fractures and natural fractures (With NF).

Uncertain parameter	Unit	Without NF (Iteration 10 Case 10)	With NF (Iteration 12 Case 7)
Matrix Permeability	md	0.0366	0.0275
Fracture Height	ft	62	53
Fracture Half-length	ft	233	257
Fracture Conductivity	md-ft	62	76
Fracture Water Saturation	-	0.74	0.55
Matrix Water Saturation	-	0.34	0.42
Exponent of krg	-	1.77	3.8
Endpoint of krw	-	0.53	0.52
Exponent of krw	-	3.85	2.5
Fracture Width	ft	0.95	0.68
Total number of NF	-	n/a	778
NF Length	ft	n/a	538
NF Conductivity	md-ft	n/a	9.9

Table 6.7: Summary of ten uncertain parameters of the best match for both cases with and without NF.

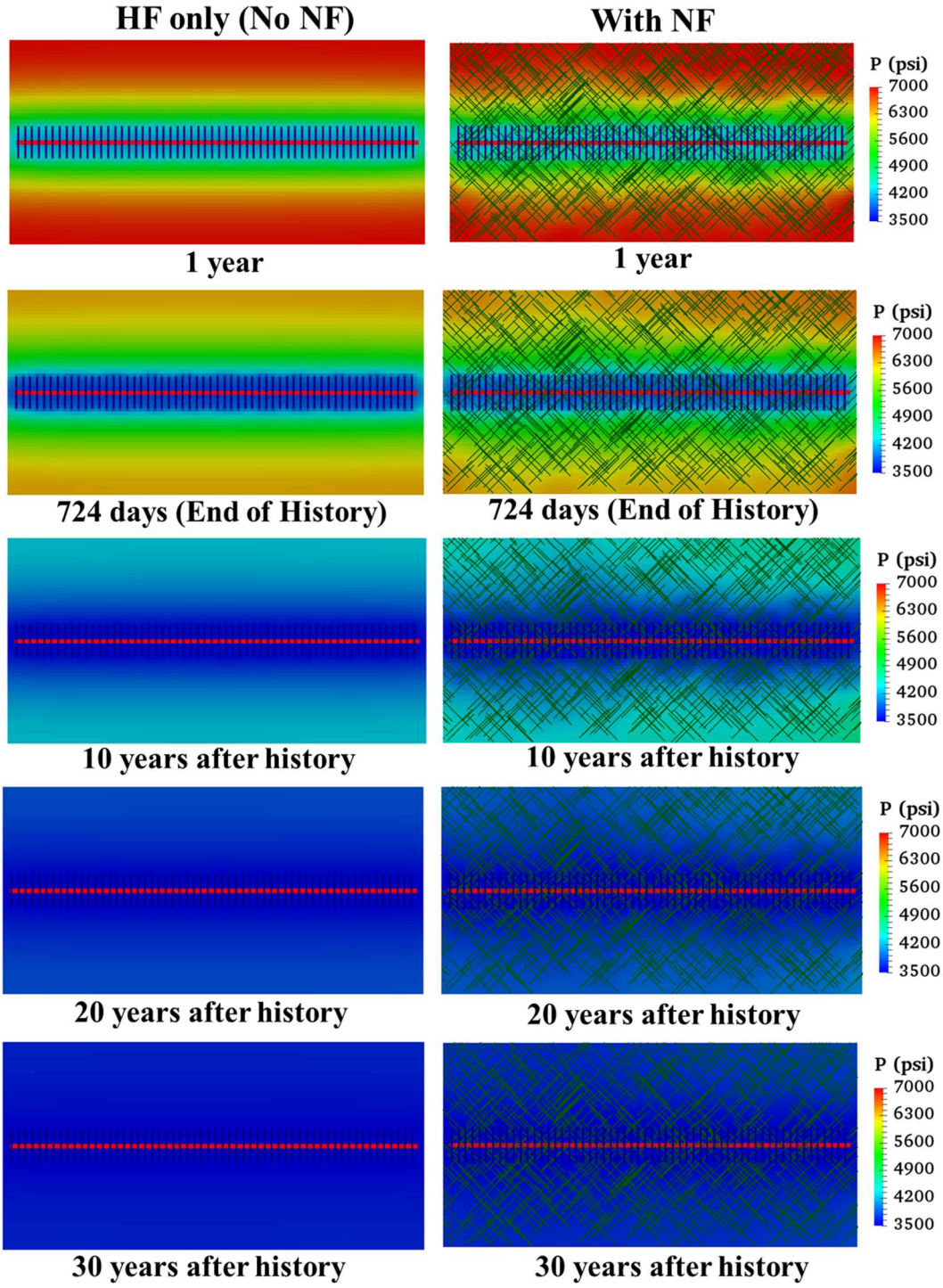


Figure 6.15: Pressure distributions comparison between case 1 with hydraulic fractures only (No NF) and case 2 with hydraulic fractures and natural fractures (With NF).

6.4 History Matching Result Discussion

In this section, we emphasize the performance of NN-MCMC algorithm in the proposed AHM workflow by presenting the history matching results and analysis of the case 2 with hydraulic fractures and natural fractures (With NF).

Figure 6.16 shows the parallel coordinate plot of history matching solutions and non-history matching solutions. Each line represents one realization of uncertain parameters. We can see that only some combinations of uncertain parameters can be history-matching solutions; for example, ones with lower fracture half-length and higher fracture conductivity in the ranges tend to be history matching solutions. Another thing to be highlighted is that NN-MCMC algorithm already explored and validated other combinations but they could not reproduce as history match solutions, as shown as the grey lines in Figure 6.16. This confirms the NN-MCMC algorithm efficiency to explore throughout the possible uncertain parameter domain.

As NN-MCMC algorithm performed automatically and proposed new cases in each iteration to be validated with reservoir simulator, Figure 6.17 shows the objective function values from simulation results vs. simulation number index. We can observe that objection function values are improving as iteration goes by (higher simulation number index) and more solutions are found at the later iteration due to the improvement of NNs. For each uncertain parameter, Figure 6.18 illustrates how the proposed cases of each uncertain parameter vs. simulation number index progresses by NN-MCMC algorithm. We can observe that the proposed points are converged to the posterior distribution as simulation number index increases especially fracture height and fracture half-length emphasizing the performance of NN-MCMC algorithm. Only certain ranges can be history-matching solutions as shown in filled red points. In addition, Figure 6.19 displays the cross plot

between proxy or NNs estimation vs. simulation results for global error, BHP error and WGR error. We can see the improvement of NN quality by comparing to the unit slope line. As iteration number is higher, the proposed points are closer to the unit slope line. Moreover, Figure 6.20 shows details of each 12 NN or proxy estimate vs simulation results. As iteration progresses, the NN prediction quality is better and closer to the unit slope line. To sum up, the NN-MCMC algorithm in the proposed AHM workflow can perform history matching automatically and find multiple history matching solutions efficiently.

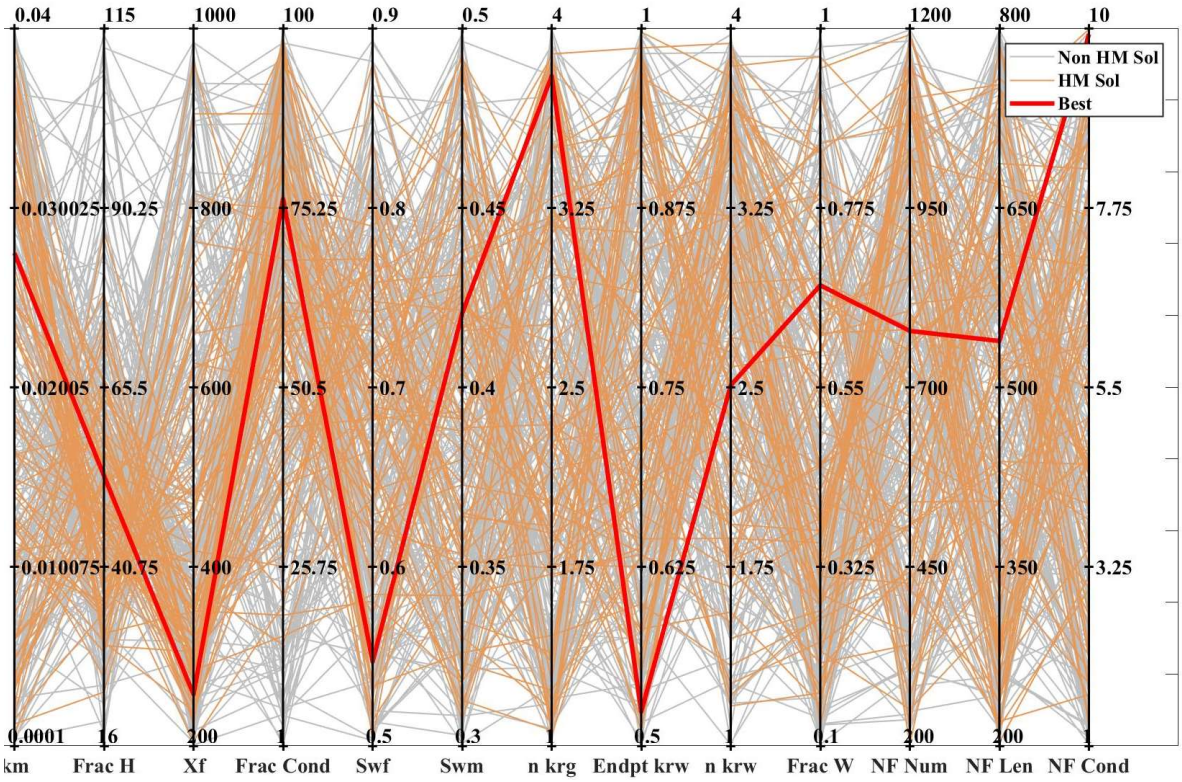


Figure 6.16: Parallel coordinate plot of all cases proposed by NN-MCMC algorithm of case 2 with hydraulic fractures and natural fractures (With NF).

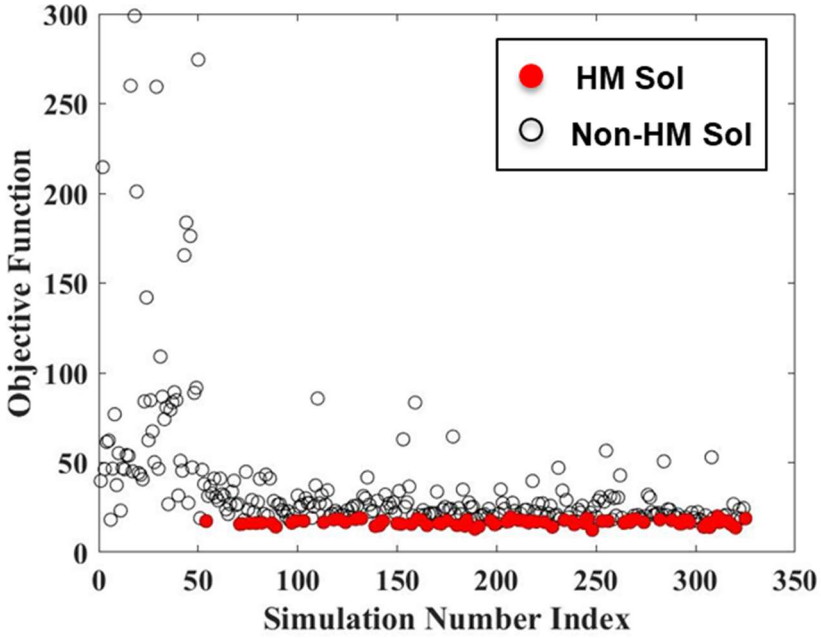
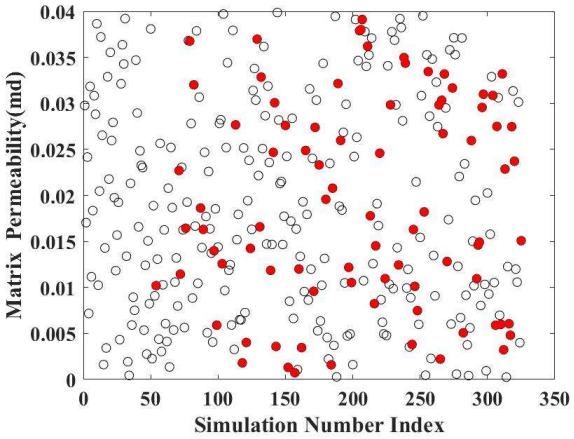
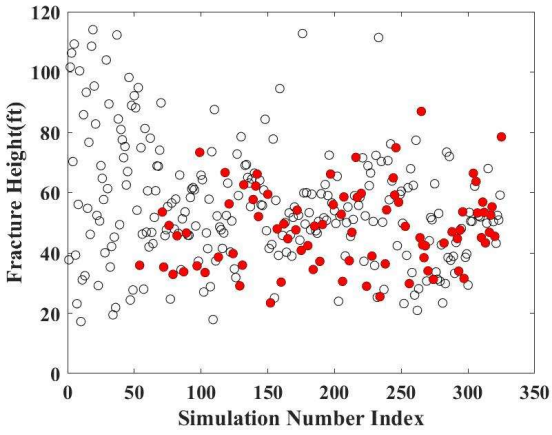


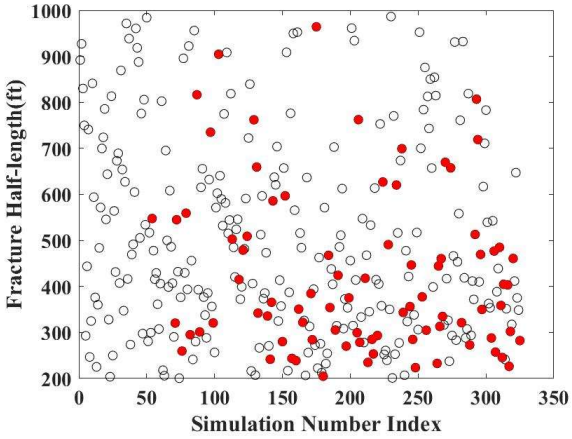
Figure 6.17: Objective function value or global error vs. simulation number index of case 2 with hydraulic fractures and natural fractures.



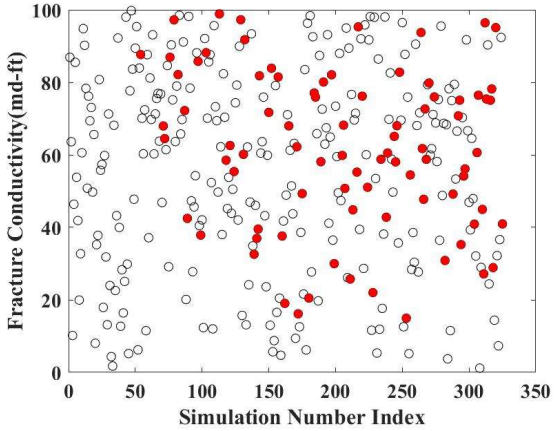
(a) Matrix permeability



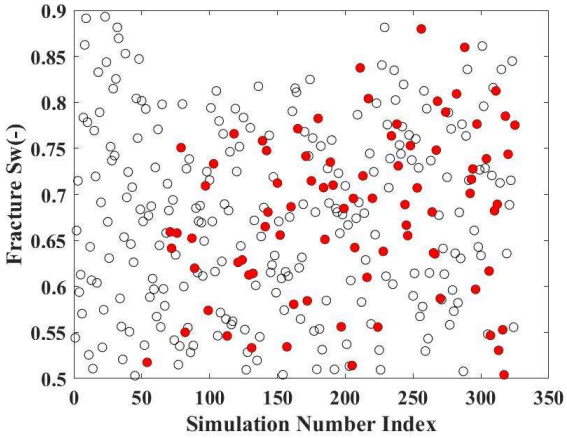
(b) Fracture height



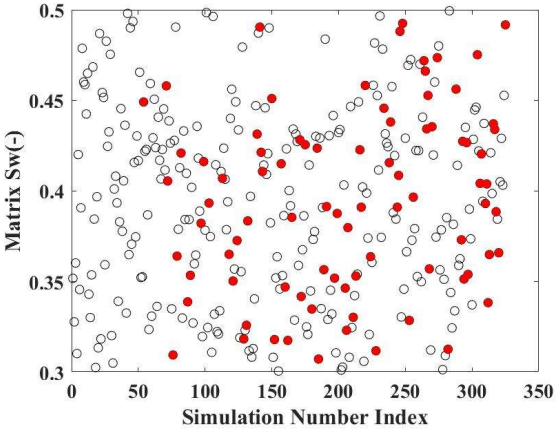
(c) Fracture half-length



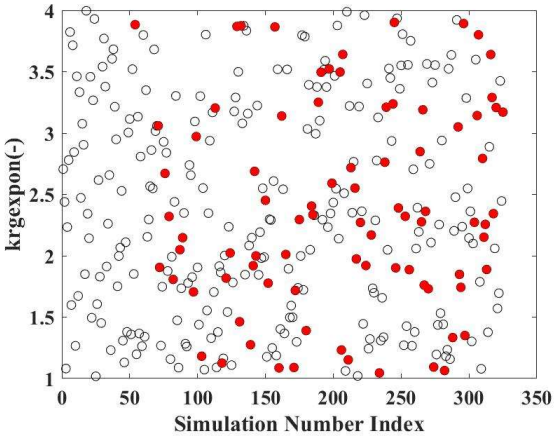
(d) Fracture conductivity



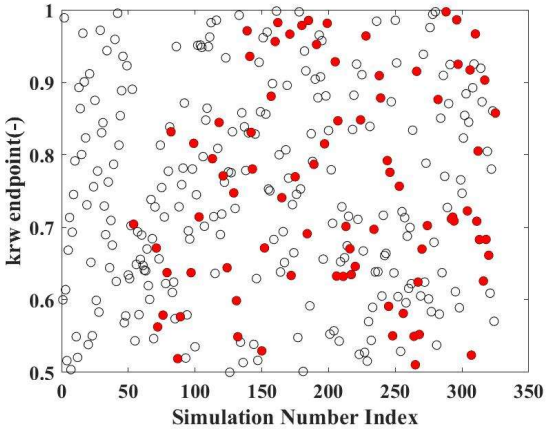
(e) Fracture water saturation



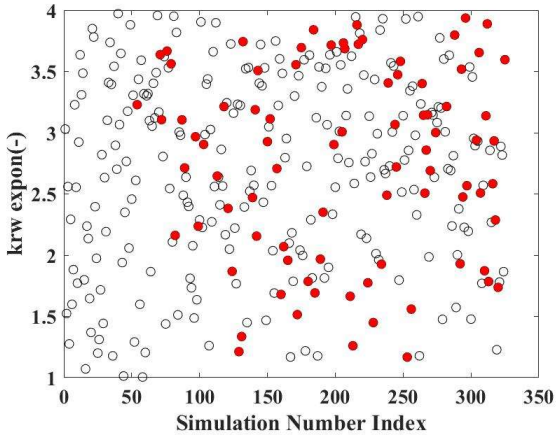
(f) Matrix water saturation



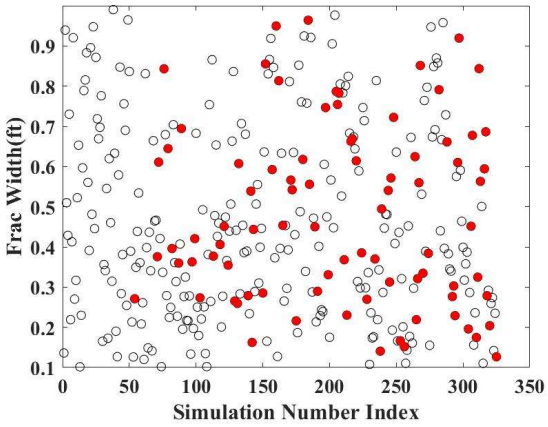
(g) Exponent of k_{rg}



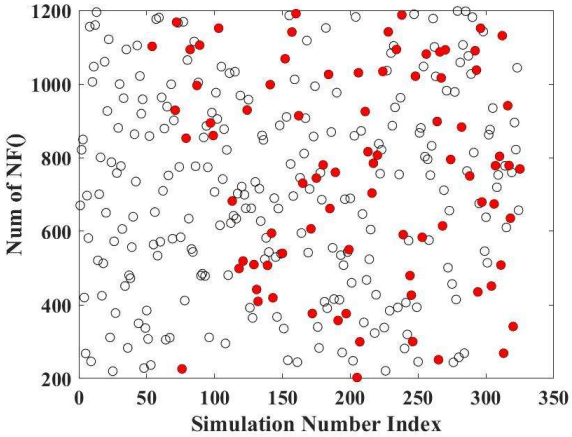
(h) Endpoint of k_{rw}



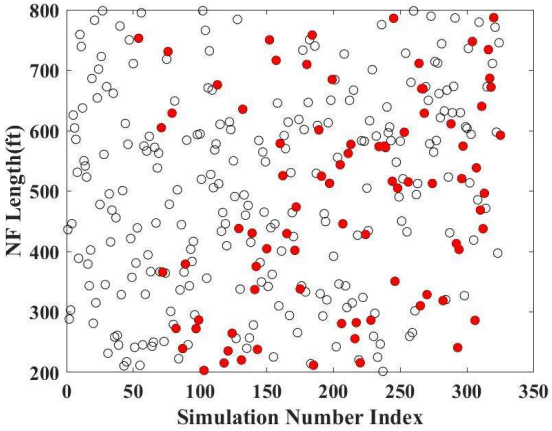
(i) Exponent of k_{rw}



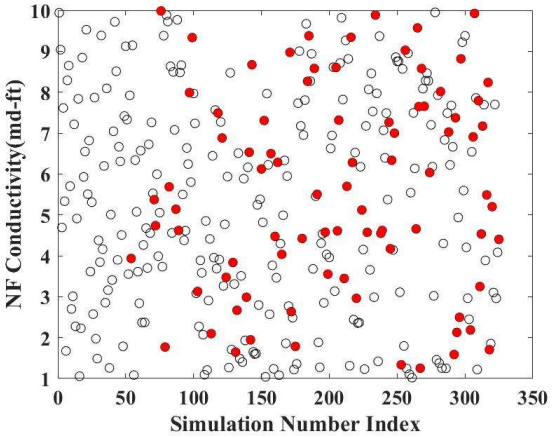
(j) Fracture width



(k) Number of NF

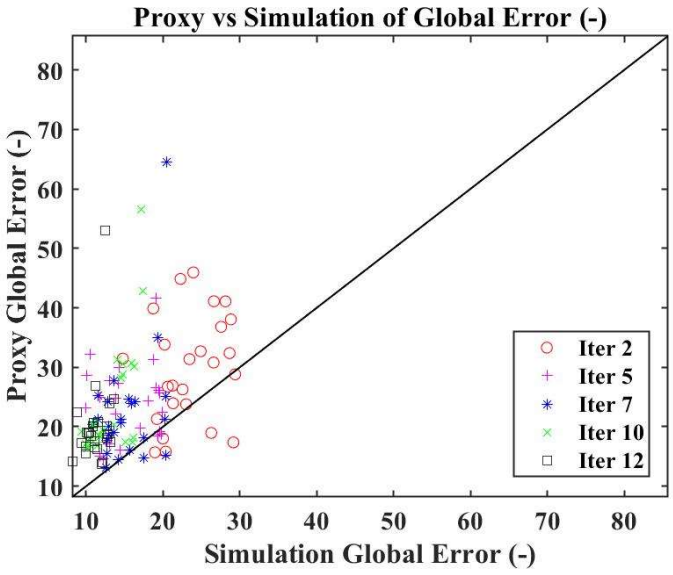


(l) NF length

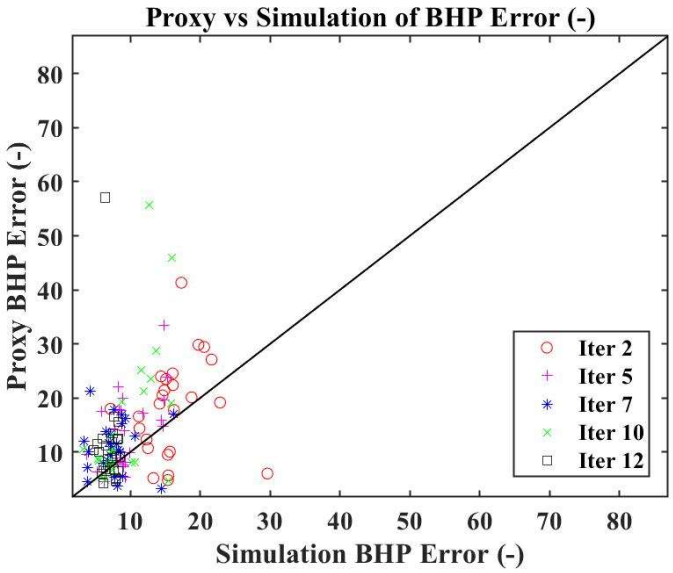


(m) NF conductivity

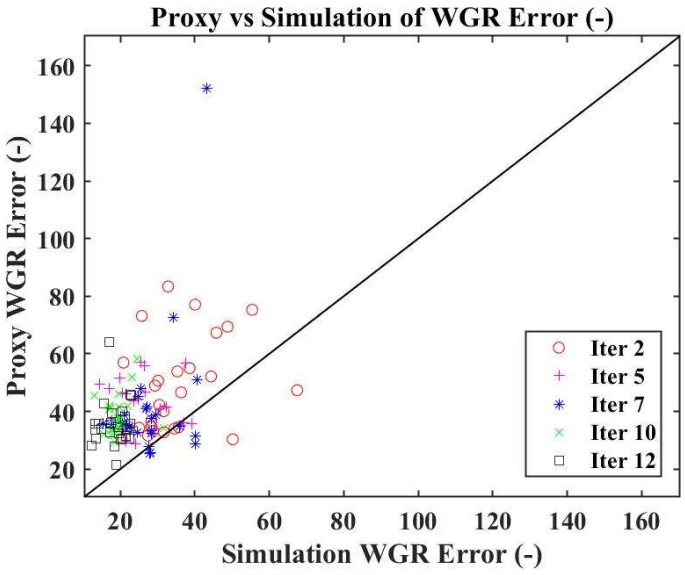
Figure 6.18: Uncertain parameter values vs. simulation number index. Black circles are non-HM solutions proposed by NN-MCMC algorithm in each iteration while filled red circles are HM solutions screened by the HM solution criteria.



(a) Global error

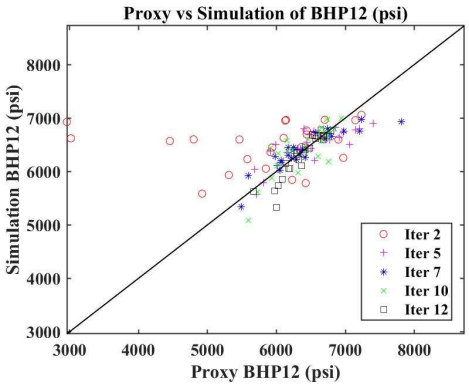


(b) BHP error

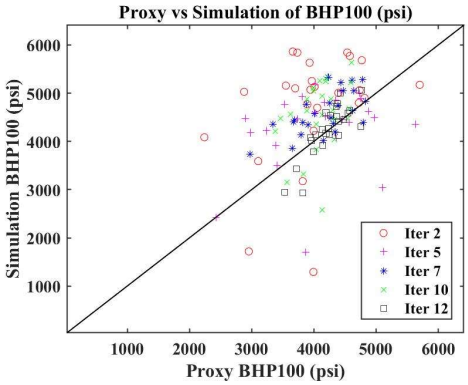


(c) WGR error

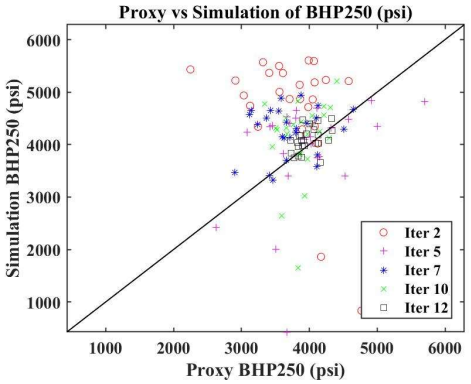
Figure 6.19: Cross plot between proxy and simulation values of (a) global error (calculated from 12 proxies), (b) BHP error (calculated from 6 BHP proxies) and (c) WGR error (calculated from 6 WGR proxies).



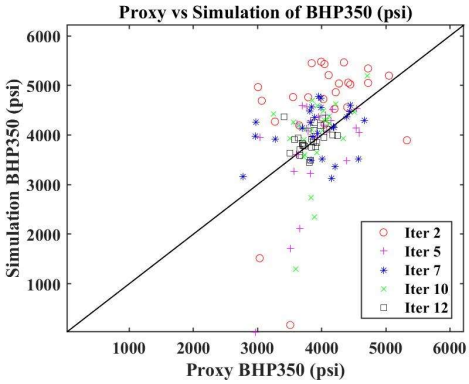
(c) BHP at day 12



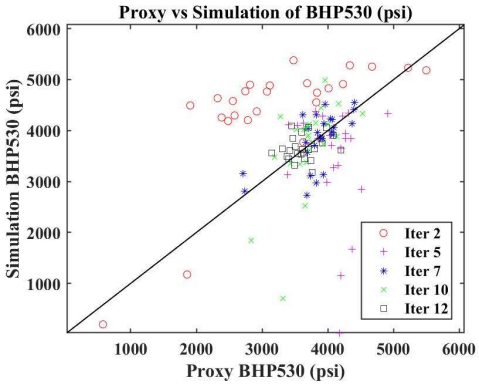
(d) BHP at day 100



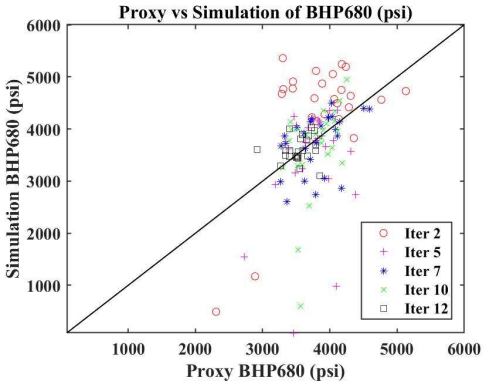
(c) BHP at day 250



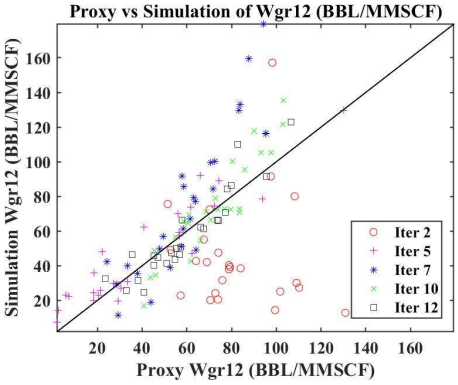
(d) BHP at day 350



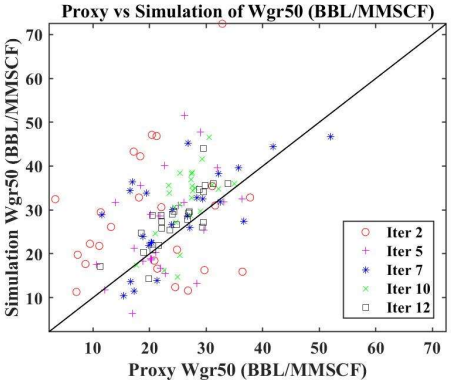
(e) BHP at day 530



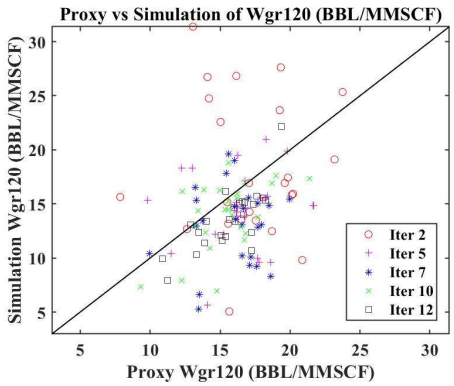
(f) BHP at day 680



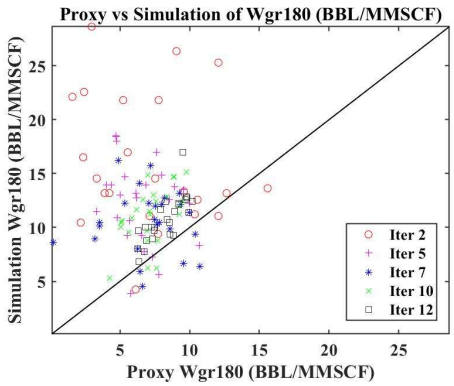
(g) WGR at day 12



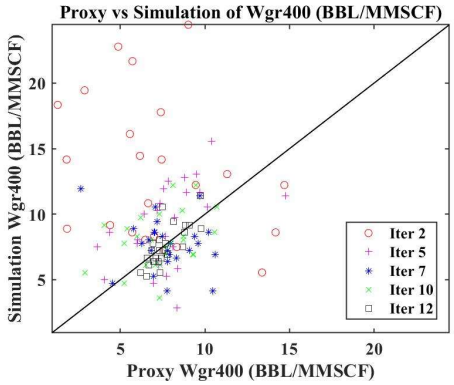
(h) WGR at day 50



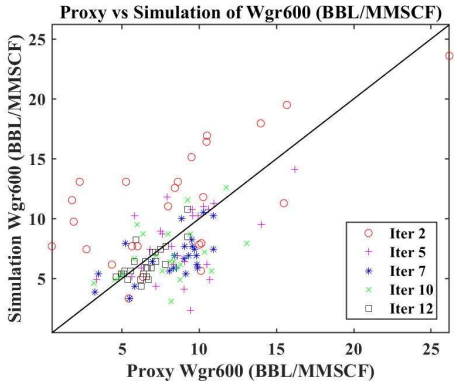
(i) WGR at day 120



(j) WGR at day 180



(k) WGR at day 400



(l) WGR at day 600

Figure 6.20: Cross plot between proxy and simulation values of 12 proxies of case 2 with hydraulic fractures and natural fractures.

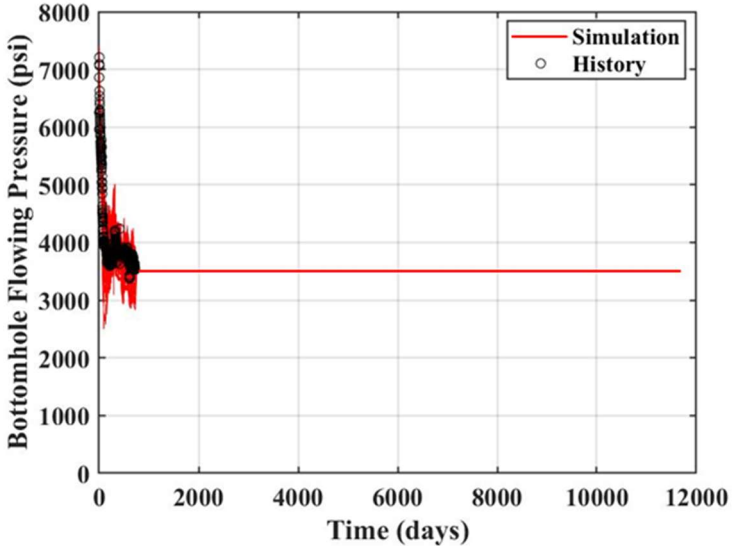
6.5 PRODUCTION FORECAST

After we obtained multiple realizations of history matching solutions from the AHM workflow, we can perform production forecast probabilistically. In this study, we forecasted the production for 30 years. In order to perform a reasonable production forecast, we assumed the constant BHP constant of 3500 psi for prediction (Figure 6.17a). The reason is that the BHP has been constant around 3500 psi for several months at the end of historical period. We performed production forecast for both cases with and without

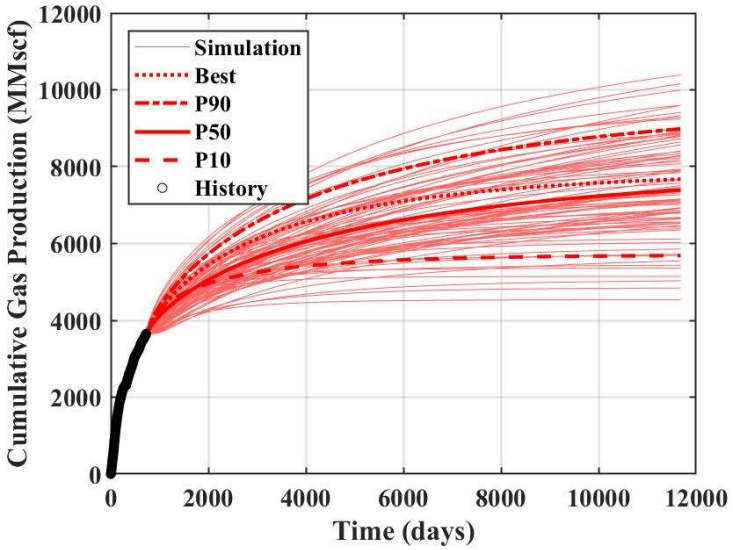
natural fractures. The simulation results of production forecast for both cases are plotted in Figure 6.21-6.22 together with the best match and P10-50-90 cases.

To compare the gas EUR between both cases, the empirical cumulative distribution function (ECDF) curves of gas EUR are shown in Figure 6.23. With the existence of natural fractures, the ECDF of gas EUR is lower than the case with only hydraulic fractures. This is because lower fracture height posterior distribution is obtained during history matching when natural fractures exist. In other words, we would over-predict the gas production if we assumed this well with only hydraulic fractures in spite the fact that there are natural fractures in the system. The degree of gas EUR over-prediction would depend on natural fractures properties and reservoir properties, and certainly, these would be different for each reservoir. In this case, we observe around 10 % discrepancies of predicted gas EUR.

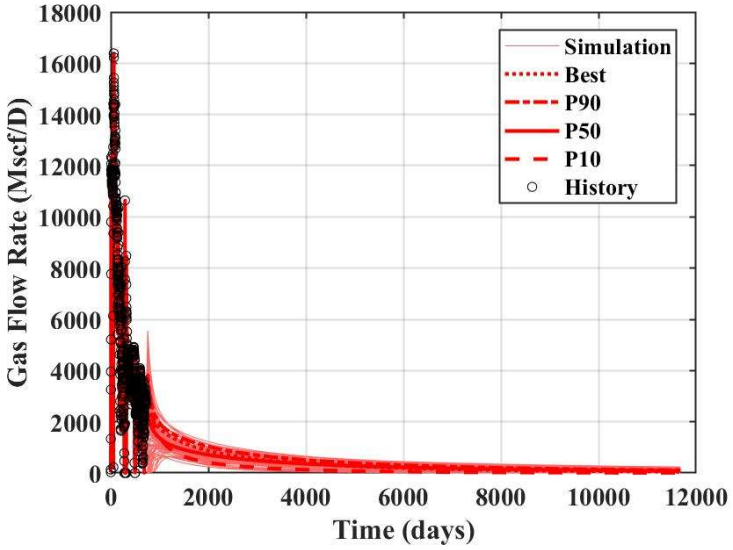
To conclude, if there are natural fractures in an actual reservoir, they should be included into reservoir model for history matching, in other words, we cannot represent the whole system of natural fractures and hydraulic fractures by a simplified model with hydraulic fractures only.



(a) BHP

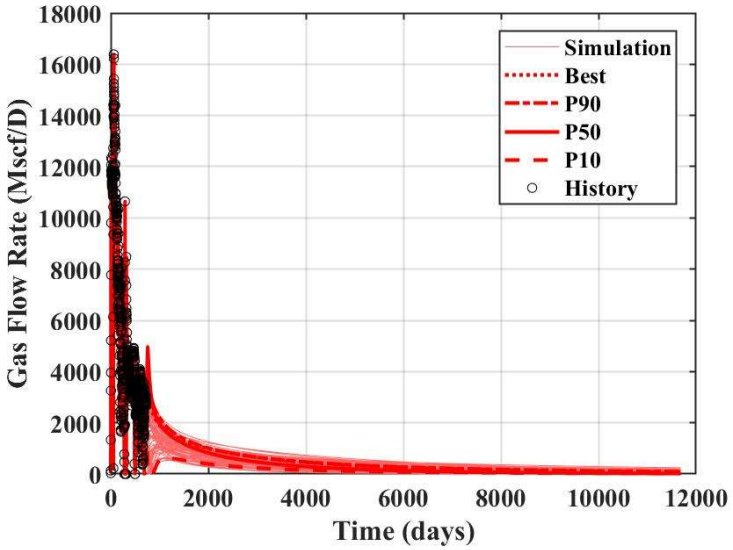


(b) Cumulative gas production

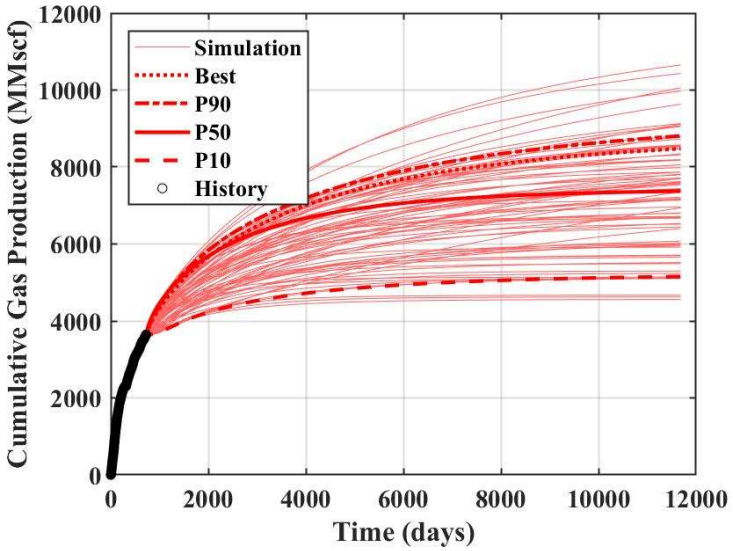


(c) Gas flow rate

Figure 6.21: Production forecast of history matching solutions of case 1 with hydraulic fractures only (No NF): (a) BHP (b) Cumulative gas production and (c) Gas flow rate.



(a) Cumulative gas production



(b) Gas flow rate

Figure 6.22: Production forecast of history matching solutions of case 2 with hydraulic fractures and natural fractures (With NF): (a) Cumulative gas production and (b) Gas flow rate.

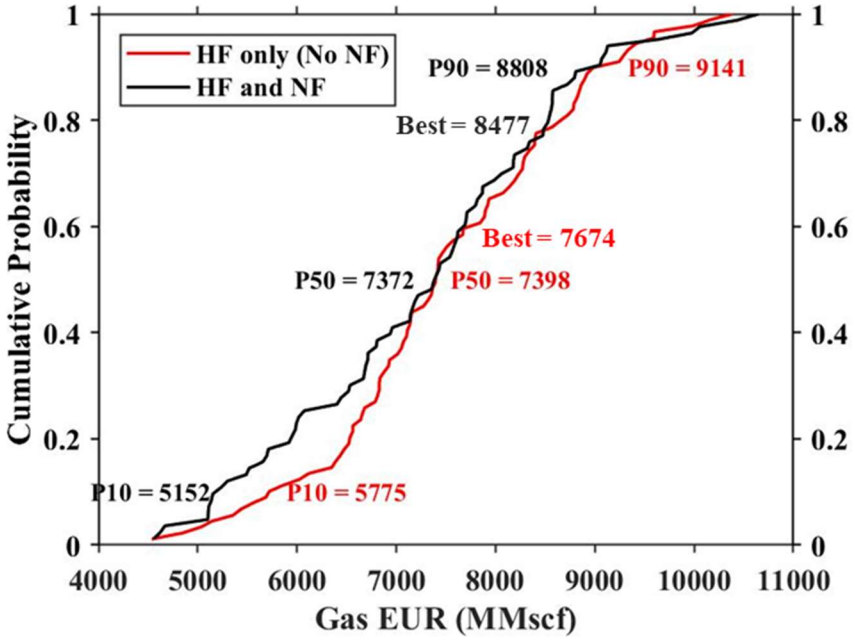
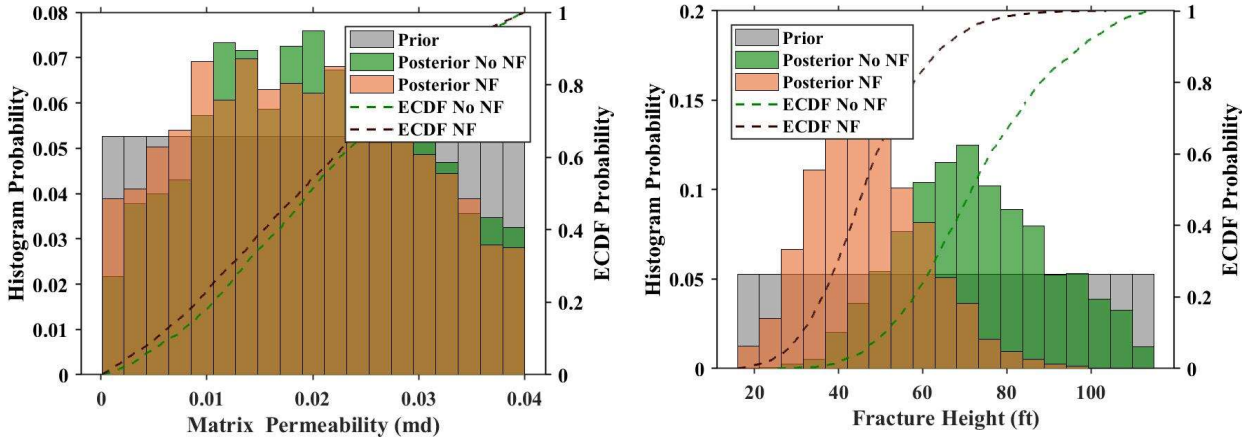


Figure 6.23: The comparison of Gas EUR empirical cumulative distribution function between case 1 with hydraulic fractures only and case 2 with hydraulic fractures and natural fractures.

6.6 1000 History Solutions from Neural Networks

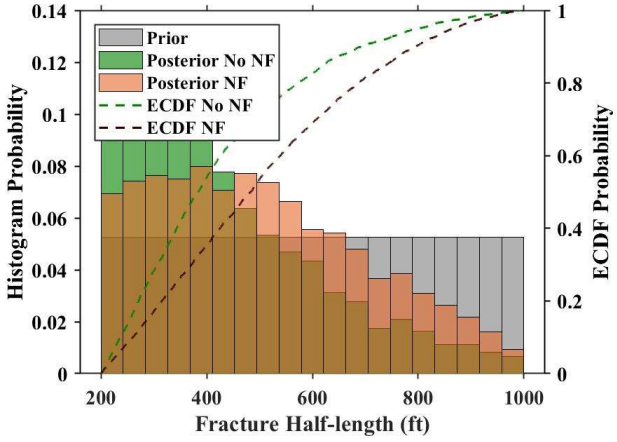
In case that we would like the smoother posterior distribution of uncertain parameters, we can use the latest proxy model or NN to generate more history matching solutions such as 1000 cases. To illustrate, Figure 6.24 shows the posterior distribution and ECDF of 1000 history matching solutions for both case 1 with hydraulic fractures only (No NF) and case 2 with hydraulic fractures and natural fractures (With NF). We can observe that the distributions are similar to those of 89 and 83 history-matching solutions of case 1 and case 2 (Figure 6.14) but with the smoother distribution. Furthermore, NNs can be created for gas EUR prediction based on the production forecast results of 89 and 83 history matching solutions of each case. Then, we can estimate gas EUR of 1000 history matching solutions from proxy (NN), as plotted in Figure 6.25. The gas EUR ECDF of the 1000

solutions from proxy yields approximately the same range compared to the actual history solutions. However, there is some slight difference between actual history matching solutions and 1000 solutions from proxy. The reasons are first, these 1000 solutions from proxy have not been validated with the actual simulator whether they are all history matching solutions or not as it is infeasible to perform and second, the oil and gas EUR proxy or NN in this study could contribute to the EUR prediction inaccuracy as we did not run all production forecast of all 1000 cases directly. Therefore, we have to use the outputs and analysis of proxy solutions with caution and should use them just to be compared to the actual history matching solutions from the AHM workflow.

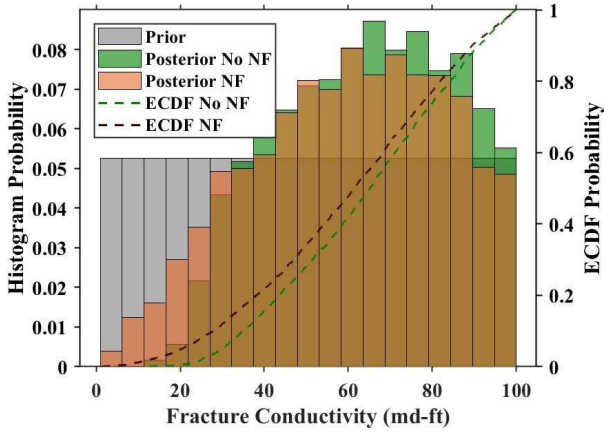


(a) Matrix permeability

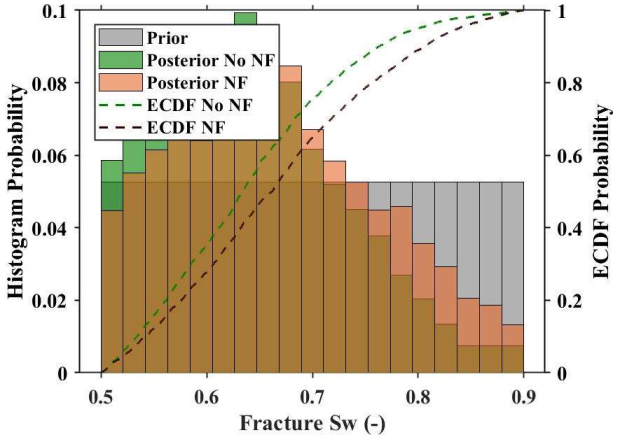
(b) Fracture height



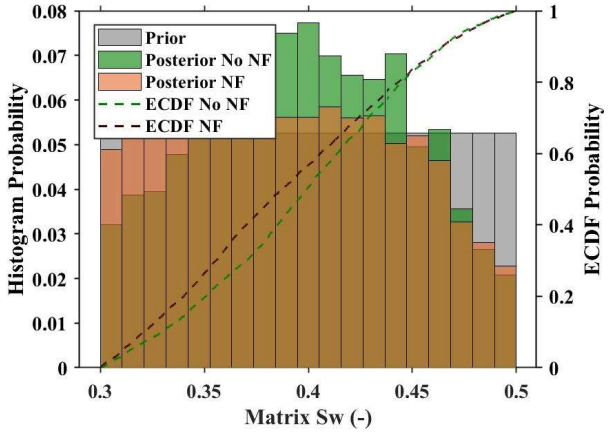
(c) Fracture half-length



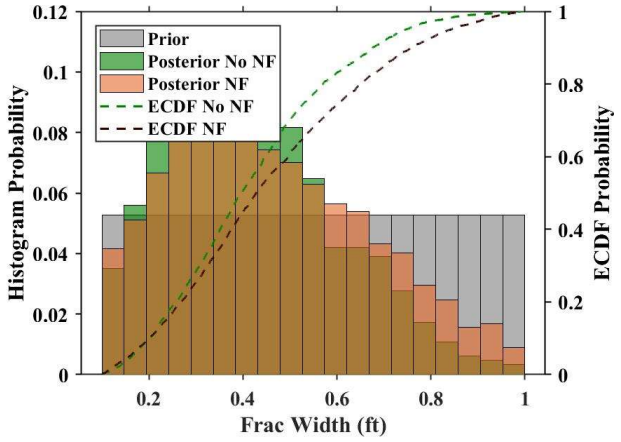
(d) Fracture conductivity



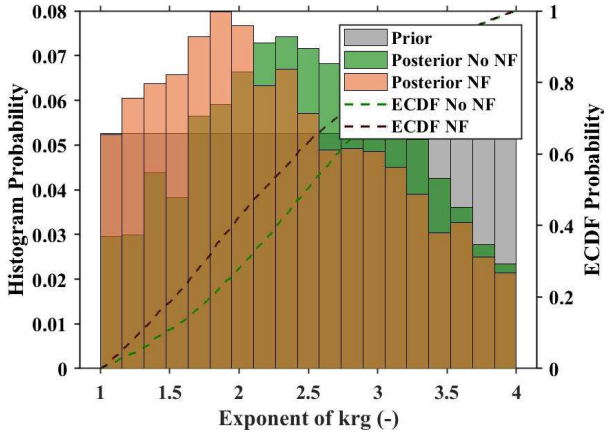
(e) Fracture water saturation



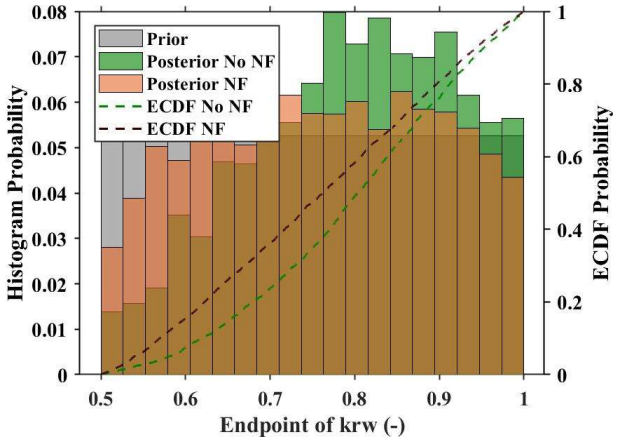
(f) Matrix water saturation



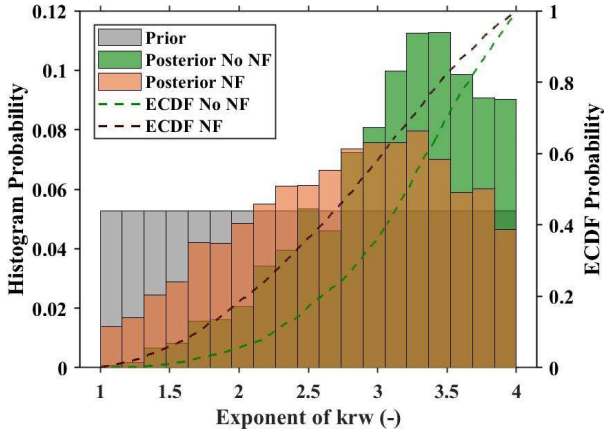
(g) Exponent of k_{rg}



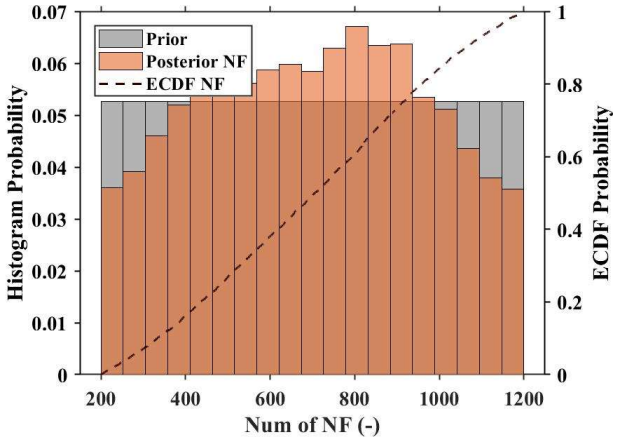
(h) Endpoint of k_{rw}



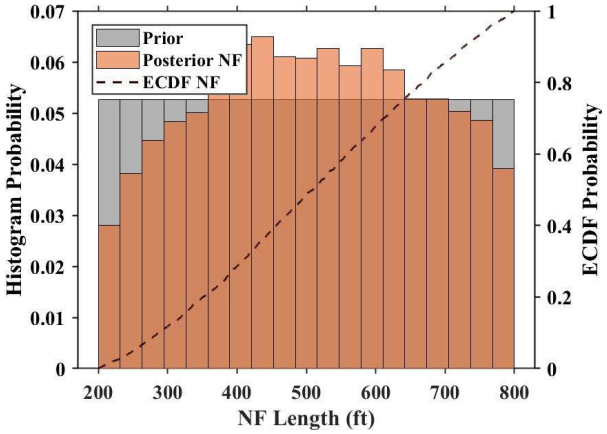
(i) Exponent of k_{rw}



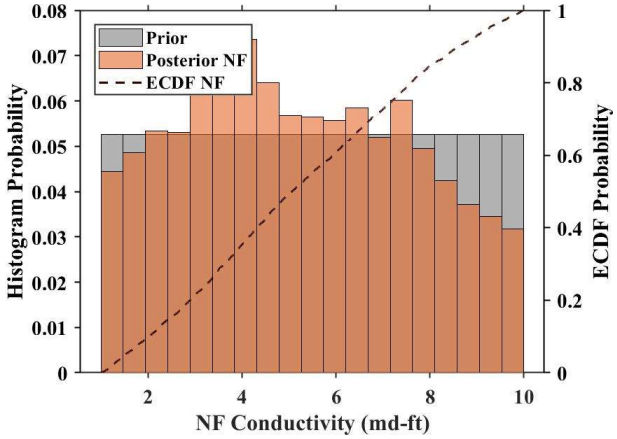
(j) Fracture width



(k) Number of NF

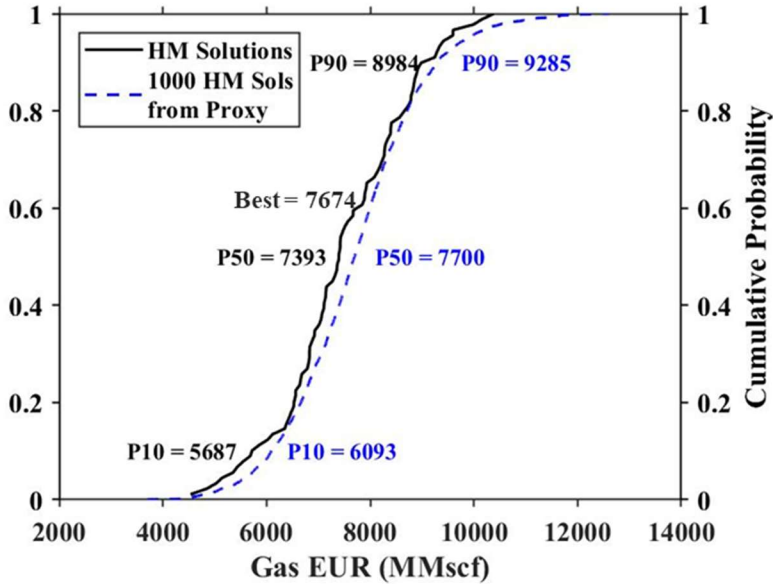


(l) NF length

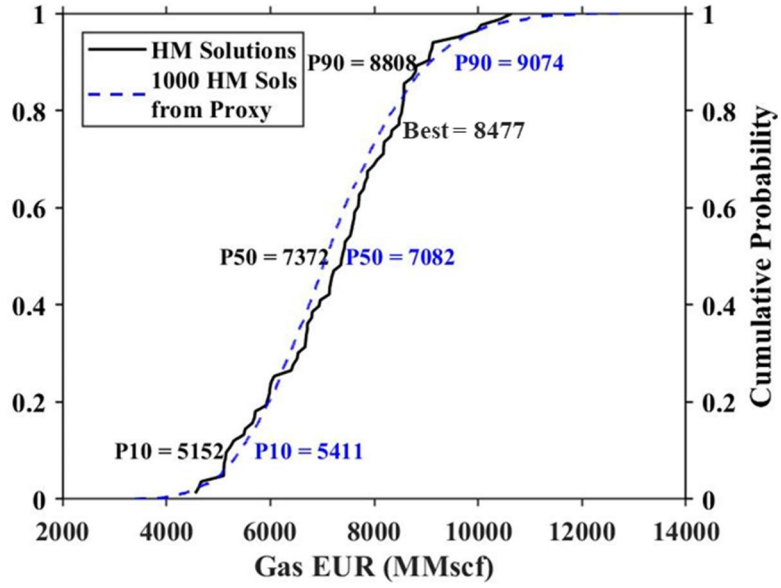


(m) NF conductivity

Figure 6.24: Prior and comparison of posterior distribution of 1000 history matching solutions from proxy between case 1 with hydraulic fractures only (No NF) and case 2 with hydraulic fractures and natural fractures (With NF).



(a) Case 1 with hydraulic fractures only (No NF)



(b) Case 2 with hydraulic fractures and f natural fractures (With NF)

Figure 6.25: The comparison of gas EUR empirical cumulative distribution function (ECDF) between history matching solutions from the AHM workflow and 1000 history matching solutions generated by proxy (NN): (a) Case 1 with hydraulic fractures only (No NF) and (b) Case 2 with hydraulic fractures and natural fractures (With NF).

6.7 BENEFITS FROM THE STUDY

The benefits from this work are that we can characterize fracture properties such as fracture height, fracture half-length and fracture conductivity and more importantly, quantify their uncertainties from production data as shown in Figure 6.9, 6.13 and 6.14. With this information, they can be integrated to calibrate fracture propagation model to improve any future hydraulic fracturing design and infill well placement optimization. For example, we can obtain dimensionless fracture conductivity (FCD) as shown in Figure 6.26 from this AHM workflow by calculating from fracture conductivity, fracture half-length and matrix permeability values of each realization. The results of FCD could be used to reflect the effectiveness of hydraulic fracturing for this well and proppant transport performance. In this case, we can interpret that the created hydraulic fractures are still finite fracture conductivity based on FCD values and can be improved.

Besides, we can characterize reservoir properties and their uncertainties such as matrix permeability, matrix water saturation and relative permeability which are sometimes challenging to be acquired by other methods for shale reservoirs. To illustrate, the relative permeability curves for shale reservoirs are sometimes infeasible to perform in the laboratory. However, with the AHM workflow, the relative permeability curves can be characterized and their uncertainty can be captured. As shown in Figure 6.27, relative permeability curves of all 83 history matching solutions together with the best match case are plotted. Furthermore, in case we can characterize those reservoir properties by other methods, the reservoir properties information from the workflow can be integrated with other techniques such as petrophysical interpretation and core data to produce more comprehensive analysis. Lastly, we can reduce the uncertainty of natural fractures from the AHM workflow and included them as the part of realization uncertainty as natural

fractures could play an important role in the long-term production forecast as demonstrated in section 6.5. Furthermore, the probabilistic production forecast, capturing the uncertainty during history matching, can be obtained from the physical-based numerical model rather than using the empirical method such as decline curve analysis alone.

6.8 CONCLUSIONS

We developed the Neural-Network Markov Chain Monte Carlo algorithm (NN-MCMC) for the assisted history matching (AHM) workflow. We utilized the proposed algorithm and AHM workflow to perform history matching with a real shale gas well automatically. Two scenarios including hydraulic fractures only (No NF) and hydraulic fractures and natural fractures (With NF) were investigated. From the AHM workflow, the effect of natural fractures existence on the posterior distribution of uncertain parameters can be evaluated. Besides, we can obtain multiple history matching solutions and characterize hydraulic fractures, natural fractures and reservoir properties in the probabilistic manner. This information can be used further for hydraulic fracturing design improvement and well placement optimization. After that, we performed production forecast for both scenarios and found that there would be some difference in gas production forecast. If we simplified the system of hydraulic fractures with natural fractures by only modeling hydraulic fractures, we would have over-predicted gas EUR.

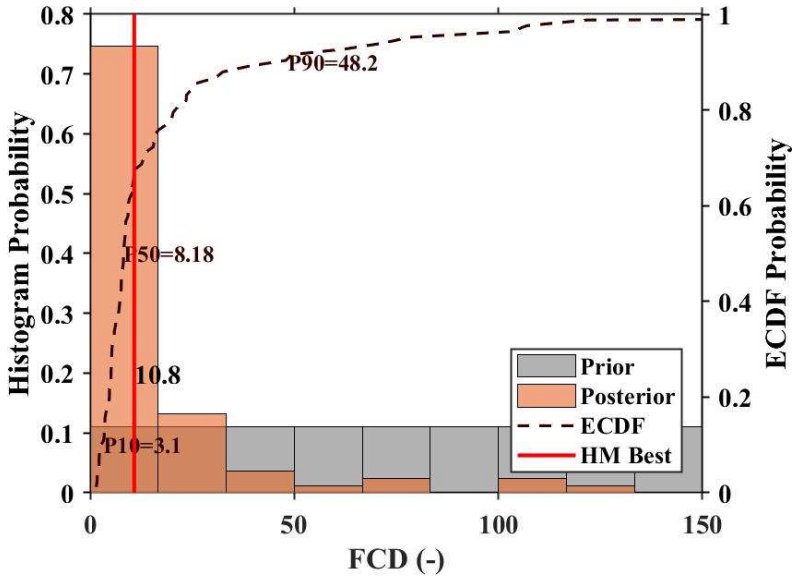


Figure 6.26: Dimensionless fracture conductivity (FCD) of 83 history matching solutions of case 2 with hydraulic fractures and natural fractures (With NF) shown with the best match case.

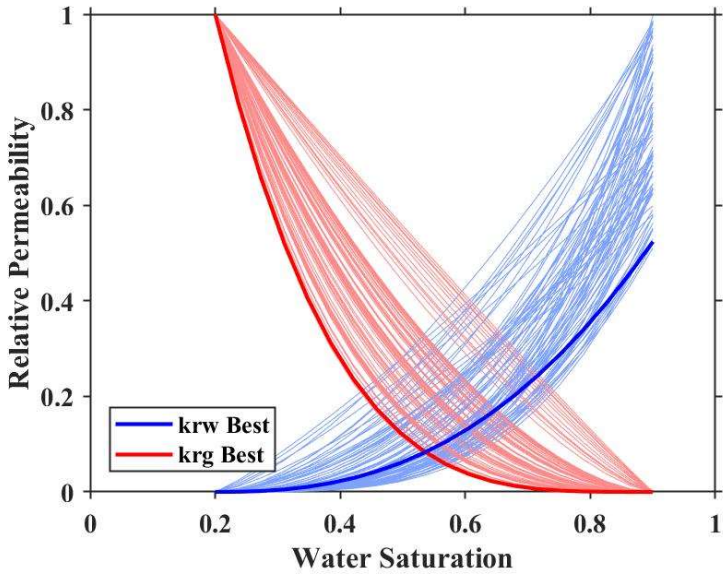


Figure 6.27: Relative permeability curves of all 83 history matching solutions of case 2 with hydraulic fractures and natural fractures (With NF) shown with the best match case.

Chapter 7: Summary, Conclusions, and Recommendations for Future Work

In this chapter, we summarize the findings from the study and also present recommendations for future work.

7.1 Summary and Conclusions

We developed the Assisted History Matching (AHM) workflow for shale reservoirs with natural fractures. We adopted a practical, efficient and accurate AHM workflow by using Embedded Discrete Fracture Model (EDFM), commercial simulator and multiple proxy-based Markov Chain Monte Carlo (MCMC) algorithm. Also, instead of retrieving one matched model, we obtained automatically multiple solutions of history matching and their posterior distributions of uncertain parameters including fractures geometry and properties, reservoir properties and natural fractures properties from the AHM workflow.

In Chapter 4, we demonstrated the application AHM workflow to a real shale gas well. Firstly, the comparison between local grid refinement (LGR) and EDFM was performed and confirmed the accuracy and efficiency of using EDFM to model fractures non-intrusively in the AHM workflow. Then, the step of parameter identification and screening was performed to determine the uncertain value ranges and to optimize number of uncertain parameters to be later used during history matching step. After that, we used proxy-based MCMC algorithm with k-nearest neighbors as proxy model to perform AHM automatically, and we then obtained the posterior distribution of uncertain parameters. Lastly, we presented probabilistic production forecast by using multiple matched models in order to account for the non-uniqueness of history matching.

Next, in Chapter 5, we applied the AHM workflow to another real shale gas well. The effects of pressure-dependent permeability and gas desorption were included into shale

gas reservoir modeling. Then, we investigated the performance of different proxy models in proxy-based MCMC algorithm including neural network and k-nearest neighbors. We found that neural network performed better in terms of accuracy and requires less number of simulations than k-nearest neighbors in order to achieve the same number of history matching solutions. However, k-nearest neighbors have benefits of requiring a lower running time by a half of one in neural networks. Therefore, which proxy model to be used depends on the limitation of each project either number of simulations or elapsed running time is more important aspect.

Then, we investigated on a more realistic concept of modeling hydraulic fractures. We compared the cases with and without enhanced permeability area (EPA) around main hydraulic fractures. By performing AHM with neural network-MCMC algorithm, we found that the existence of EPA would affect the history matching solutions. The posterior distributions of fracture geometries and properties and matrix permeability shifted to the lower values. These changes in history-matched solutions have an impact on probabilistic production forecast. In this case study, we observed that the case with EPA has the lower gas estimated ultimate recovery (EUR) around 25-30% compared with the case without EPA. This means if we simplified the simulated rock volume created during hydraulic fracturing by modeling only hydraulic fractures, this could over-estimate fracture geometry and gas EUR prediction.

Lastly, in Chapter 6, we continued applying the AHM workflow to a shale gas well to investigate the impact of natural fractures existence. The effects of gas desorption and pressure-dependent permeability were also included in shale gas reservoir modeling. Then, we performed two cases with and without natural fractures as uncertain parameters adopting neural network-MCMC algorithm. Similarly, we examined the effects of natural fractures on the history matching solutions and production forecast. In this case, we found

that the case with natural fractures has posterior distributions significantly lower in fracture height, lower in fracture conductivities and higher in fracture water saturation than the case without natural fractures. Also, we observed less curvature on relative permeability curves of krg and krw. All changes due to natural fractures existence are to compensate more flow contributed by natural fractures and changes in other uncertain parameters. For production forecast, in this case, we found that the case with natural fractures has lower gas EUR around 10% compared with the case without natural fractures due to lower fracture height. With this effect on production forecast, this means a decision to include natural fractures modeling in the actual system of natural fractures and hydraulic fractures will depend on the natural fractures and reservoir properties whether they have a significant effect on history matching solutions and production forecast or not. However, these effects can only be quantified after we included natural fractures as uncertain parameters in AHM.

One of benefits of this study is that we can characterize fracture geometries and properties, reservoir properties and natural fracture uncertainties in a probabilistic manner from the production data which is always available at no additional cost. These information from history matching can be used to calibrate fracture propagation model to optimize any future hydraulic fracturing design, infill well placement decision. Lastly, we can forecast production in a probabilistic way to capture subsurface uncertainties and the non-uniqueness of history matching problem.

7.2 Recommendations for Future Work

Proxy-based MCMC Algorithm and AHM Workflow Codes

- In this study, we used only one long MCMC chain during proxy-based MCMC algorithm. The effect of using multiple short MCMC chains has not been studied.

Therefore, the investigation of using multiple short MCMC chains vs. one long MCMC chain could determine whether there is any significant difference or not and which options gives a better efficiency of proxy-based MCMC algorithm.

- The current study handled the convergence of MCMC by using a high number of MCMC chain. However, the further development of evaluating MCMC chain convergence specifically for each case rather than one constant number of MCMC steps could improve the efficiency of the algorithm.
- The number of initial points in the first iteration in this study is based on experience and empirical method, which is five times the number of uncertain parameters. However, the study on varying number of initial points could help determine the optimum number of initial points and could improve the efficiency and accuracy of the proxy-based MCMC algorithm
- The current study only adopted Metropolis-Hasting (MH) MCMC, which is a random walk with probability. Future work could evaluate the performance comparison between MH MCMC and other types of MCMC such as Gibbs and Hamiltonian MCMC in terms of the number of history matching solutions found and computational efforts.
- In chapter 5, we compared the performance of neural network vs. k-nearest neighbors and found that neural network provides a lot more accuracy in terms of number of history matching solutions found but with higher elapsed time. Therefore, to speed up neural network related calculation, the study to find alternative ways of training neural network could be beneficial to the history matching algorithm in order to achieve approximately the same accuracy with less computational time. Also, the architecture of neural network we used in this study is fully connected neural network and has one hidden layer of 10 neurons. While many literatures recommend focusing on the quality

of training data rather than finding the optimized neuron network architecture, another study on different architecture of neural network such as partially connected structure, number of hidden layers and number of hidden neurons can be performed whether they help optimize the history matching algorithm or not.

- The current AHM workflow codes now only linked to one commercial simulator, which is CMG such as CMG-GEM and CMG-IMEX. Extending the scope of AHM workflow codes to other commercial reservoir simulators such as ECLIPSE or UT in-house simulators could make this AHM tool to be more powerful.
- All of AHM workflow codes are currently written in MATLAB and Excel VBA. Translating the codes into more efficient and robust programming language such as Python or C++ could improve the running time of AHM workflow codes. Moreover, for the current usage of AHM workflow, users must learn and understand the source codes before they can perform AHM workflow. Therefore, the development of graphic user interface of AHM workflow can help increase the accessibility of this workflow to general users.

Fractures Modeling

- The current study modeled hydraulic fractures either by planar fractures or planar fractures with enhanced permeability representing simulated rock volume (SRV). The comparison study of other SRV modeling techniques such as DFN model around main hydraulic fractures with the current methods could be useful to determine whether there is any necessity to model fractures with more complicated concept or not and may provide more realistic fracture flow behavior.
- The study to integrate the current AHM workflow with fractures output from fractures propagation model can provide more realistic and complex fractures geometry. In

addition, more constraints for history matching from other diagnostic tools and analysis such as microseismics, fiber optic distributed acoustic sensing (DAS)/distributed temperature sensing (DTS) data and Diagnostic Fracture Injection Test (DFIT) analysis can help narrowing down the parameter uncertainty ranges during AHM workflow.

Applications of AHM Workflow

- All current studies were the single well model. The study of multi-well history matching can be useful to characterize the multi-well performance in more holistic view and to account for their interaction and inference between wells.
- In the current study, the geological and petrophysical inputs were not available to provide a static model that represents heterogeneity of shale reservoirs, only homogeneous model was assumed to perform AHM. Therefore, the study to integrate the AHM workflow with multi-disciplinary inputs to obtain realistic static model accounting for shale heterogeneity can extend and illustrate the benefits of AHM workflow in a more comprehensive point of view.
- Another interesting study is to investigate the effect of different days of historical production data to the history matching solutions, posterior distribution of uncertain parameters and production forecast. The findings from the study may help determine the results validity of history matching and what concerns to be aware of when using history-matching results at different stages of well life.
- Another useful application of AHM workflow is to investigate the wells with the same completion and fracturing design but having different production performances. By performing history matching, we may find the cause of the discrepancy and solve the poor performance in low production wells.

Glossary

ACRONYMS

AHM	=	Assisted history matching
BHP	=	Bottomhole pressure
DAS	=	Distributed acoustic sensing
DFIT	=	Diagnostic fracture injection test
DFM	=	Discrete fracture model
DPDK	=	Dual porosity, dual permeability model
DTS	=	Distributed temperature sensing
ECDF	=	Empirical cumulative distribution function
EDFM	=	Embedded discrete fracture model
EPA	=	Enhanced permeability area
EnKF	=	Ensemble Kalman Filter
EnS	=	Ensemble smoother
EUR	=	Estimated ultimate recovery
FCD	=	Dimensionless fracture conductivity
HF	=	Hydraulic fractures
HM	=	History matching
ILU	=	Incomplete lower triangular upper triangular factorization
KNN	=	K-Nearest Neighbors
LGR	=	Local grid refinement
LHS	=	Latin hypercube sampling
LS-LR-DK	=	Logarithmically spaced, locally refined, dual permeability grid
MCMC	=	Markov-chain Monte Carlo
MH	=	Metropolis-Hasting

NF	=	Natural fractures
NN	=	Neural network
NNCs	=	Non-neighboring connections
PDF	=	Probability density function
PI	=	Productivity Index
RTA	=	Rate transient analysis
SRV	=	Simulated rock volume
WGR	=	Water gas ratio
WHP	=	Wellhead pressure

NOMENCLATURE

A	=	Area
A_{NNC}	=	Contact area between the NNC pair
B_f	=	Formation volume factor of phase f
d_{NNC}	=	Distance between the NNC pair
g	=	Gravitational acceleration
H_s	=	Fracture-segment height
K	=	Number of K nearest points
k	=	Pressure-dependent permeability
k_f	=	Fracture permeability
k_i	=	Permeability at initial reservoir pressure
k_{NNC}	=	Permeability associated with the connection
k_{rf}	=	Relative permeability of phase f

k_s	=	Permeability
L	=	Fracture-segment length
m	=	Number of response parameters
n	=	Number of data points
NF_j	=	Difference between the minimum and the maximum value of response parameter j
P_{cgo}	=	Capillary pressure between gas and oil
P_{cow}	=	Capillary pressure between oil and water
P_f	=	Pressure of phase f
$p(\theta \mid x)$	=	Posterior distribution of the uncertain parameters given the measured production history
q_{fsc}	=	Flow rate of phase f at standard condition
r_e	=	Effective radius
r_w	=	Wellbore radius
s_f	=	Saturation of phase f
S_{seg}	=	Area of the fracture segment perpendicular to the fracture aperture
T_{NNC}	=	NNC transmissibility factor
v_b	=	Bulk volume
V_b	=	Bulk volume of fracture cell
$\vec{v}_f$	=	Velocity vector of phase f
w_f	=	Fracture width
w_{ij}	=	Weight of data points i of response parameter j
WI_f	=	Well index
X_f	=	Fracture half-length

$x_{ij,\text{model}}$ = Value from model at index *i* of response parameter *j*

$x_{ij,\text{history}}$ = Historical data at index *i* of response parameter *j*

$\hat{z}(\vec{\theta}_0)$ = Estimated response parameters of unmeasured points

$z(\vec{\theta}_i)$ = Response of measured points

α^* = Acceptance probability ratio

γ = Permeability modulus

ΔP = Potential difference

δ = Walking distance

ε^2 = Global error of current-step uncertain parameters

ε^{*2} = Global error of proposed uncertain parameters

θ = Uncertain parameter realization at the current step

θ^* = Proposed uncertain parameters realization

$\vec{\theta}_0$ = A vector of independent uncertain parameters

$\vec{\theta}_i$ = A vector of the nearest k measured points

λ = Box-Cox plot lambda

λ_l = Relative mobility of phase l

λ_{i0} = Weight function between $\vec{\theta}_0$ and $\vec{\theta}_i$

μ_f = Viscosity of phase f

ρ_f = Density of phase f

Φ = Potential equal to P_f - $\rho_f g$

$\emptyset$ = Porosity

ϕ_f	=	Effective porosity of fracture cell
∇P	=	Gradient of pressure

SUBSCRIPTS

f	=	Phase f
f, frac	=	Fracture
g	=	Gas
i	=	Index of data points
j	=	Index of response parameters
m	=	Matrix
NNC	=	Non-neighboring connection
o	=	Oil
sc	=	Standard condition
srv	=	Simulated rock volume
w	=	Water

References

Aderibigbe, A., Valdes, C., and Heidari, Z., 2016. Integrated Rock Classification in the Wolfcamp Shale based on Reservoir Quality and Anisotropic Stress Profile Estimated from Well Logs. *Interpretation* **4** (2): 1-18.

Agrawal, S., Shrivastava, K., Sharma, M., 2019. Effect of Shear Slippage on the Interaction of Hydraulic Fractures with Natural Fractures. Paper SPE 194361, presented at the SPE Hydraulic Fracturing Technology Conference and Exhibition, The Woodlands, Texas, 5-7 February.

Amin, S., Wehner, M., Heidari, Z., and Tice, M., 2016. Rock Classification in the Eagle Ford Shale through Integration of Petrophysical, Geological, and Geochemical Characterization. Paper SPWLA-2016-KKK, presented at the SPWLA 57th Annual Logging Symposium, Reykjavik, Iceland, 25-29 June.

Aziz, K., and Settari, A. 1979. *Petroleum Reservoir Simulation*. Applied Science Publishers.

Box, G.E.P. and Cox, D.R., 1964. An Analysis of Transformations, *Journal of the Royal Statistical Society*, Series B, **26** (2), 211-252.

Brown, M., Ozkan, E., Raghavan, R., and Kazemi, H., 2009. Practical Solutions for Pressure Transient Responses of Fractured Horizontal Wells in Unconventional Reservoirs. Paper SPE 125043, presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, 4-7 October.

Cavalcante Filho, J.S.A., Shakiba, M., Moinfar, A., and Sepehrnoori, K., 2015a. Implementation of a Preprocessor for Embedded Discrete Fracture Modeling in an IMPEC Compositional Reservoir Simulator. Paper SPE 173289, presented at the SPE Reservoir Simulation Symposium, Houston, Texas, 23-25 February.

Cavalcante Filho, J. S. de A., Xu, Y., and Sepehrnoori, K., 2015b. Modeling Fishbones Using the Embedded Discrete Fracture Model Formulation: Sensitivity Analysis and History Matching. Paper SPE 175124, presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 28-30 September.

Chai, Z., Tang, H., He, Y., Killough J., and Wang Y., 2018. Uncertainty Quantification of the Fracture Network with a Novel Fracture Reservoir Forward Model. Paper SPE 191395, presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas, 24-26 February.

Chang, H., and Zhang, D., 2018. History Matching of Stimulated Reservoir Volume of Shale-Gas Reservoirs Using an Iterative Ensemble Smoother. *SPE Journal* **23** (2): 346-366

Chen, C., Gao, G., Li, R., Cao, R., Chen, T., Vink, J. C., and Gelderblom, P., 2017. Integration of Distributed Gauss-Newton with Randomized Maximum Likelihood

Method for Uncertainty Quantification of Reservoir Performance. Paper SPE 182639, presented at the SPE Reservoir Simulation Conference, Montgomery, Texas, 20-22 February

Christie, M., Eydinov, D., Demyanov, V., Talbot, J., Arnold, D., and Shelkov, V., 2013. Use of Multi-Objective Algorithms in History Matching of a Real Field. Paper SPE 163580, presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, 18-20 February.

Cullick, A. S., Johnson, D., and Shi, G., 2006. Improved and More-Rapid History Matching With a Nonlinear Proxy and Global Optimization. Paper SPE 101933, presented at SPE Annual Technical Conference and Exhibition, San Antonio, Texas, 24-27 September.

Cipolla, C. L., Lolon, E., Erdle, J. C., and Rubin, B., 2009. Reservoir Modeling in Shale-Gas Reservoirs. Paper SPE 125530, presented at the SPE Eastern Regional Meeting, Charleston, West Virginia, 23-25 September.

Cipolla, C., Maxwell, S., and Mack, M., 2012. Engineering Guide to the Application of Microseismic Interpretations. Paper SPE 152165, presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, 6-8 February.

Cipolla, C., and Wallace J., 2014. Stimulated Rerservoir Volume: A Misapplied Concept?. Paper SPE 168596, presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, Texas, 4-6 February.

CMG-CMOST., 2017. CMOST User's Guide, Computer Modeling Group Ltd.

Dachanuwattana, S., Jin, J., Zuloaga-Molero, P., Li, X., Xu, Y., Sepehrnoori, K., Yu, W., and Miao, J. 2018a. Application of Proxy-based MCMC and EDFM to History Match a Vaca Muerta Shale Oil Well. *Fuel* **220**: 490-502.

Dachanuwattana, S., Xia, Z., Yu, W., Qu, L., Wang, P., Liu, W., Miao, J., and Sepehrnoori, K. 2018b. Application of Proxy-based MCMC and EDFM to History Match a Shale Gas Condensate Well. *Journal of Petroleum Science and Engineering* **167**: 486-497.

Dachanuwattana, S., Yu, W., and Sepehrnoori, K., 2019. An efficient MCMC history matching workflow using fir-for-purpose proxies applied in unconventional oil reservoirs. *Journal of Petroleum Science and Engineering* **176**: 381-395.

Ding, D., 2011. Development of a Data-Partition Technique for Gradient-Based Optimization Methods in History Matching. *SPE Journal* **16** (3):582-593.

Du, C., Zhang, X., Zhan, L., Gu, H., Hay, B., Tushingham, K., and Zeema Y., 2010. Modeling Hydraulic Fracturing Induced Fracture Networks in Shale Gas Reservoirs as a Dual Porosity System. Paper SPE 132180, presented at the CPS/SPE International Oil & Gas Conference and Exhibition, Beijing, China, 8-10 June.

Elsakout, D., Christie, M., and Lord, G., 2015. Multilevel Markov Chain Monte Carlo (MLMCMC) For Uncertainty Quantification. Paper SPE 175870, presented at the SPE North Africa Technical Conference and Exhibition, Cairo, Egypt, 14-16 September.

Emerick, A. A., and Reynolds, A. C., 2011. History Matching a Field Case Using the Ensemble Kalman Filter With Covariance Localization. *SPE Reservoir Evaluation & Engineering* **14** (4): 443-452.

Emerick, A. A., and Reynolds, A. C., 2012. Combining the Ensemble Kalman Filter With Markov-Chain Monte Carlo for Improved History Matching and Uncertainty Characterization. *SPE Journal* **17** (2): 418-440

Erbas, D., and Christie, M. A., 2007. Effect of Sampling Strategies on Prediction Uncertainty Estimation. Society of Petroleum Engineers. Paper SPE 106229, presented at the SPE Reservoir Simulation Symposium, Houston, Texas, 26-28 February.

Eltahan, E., Yu, W., Sepehrnoori, K., Kerr, E., Miao, J., and Ambrose, R. 2019. Modeling Naturally and Hydraulically Fractured Reservoirs with Artificial Intelligence and Assisted History Matching Methods Using Physics-Based Simulators. Paper SPE 195269, presented at SPE Western Regional Meeting, San Jose, California, 23-26 April.

Fiallos Torres, M.X., Yu, W., Ganjdanesh, R., Kerr, E., Sepehrnoori, K., Miao, J., Ambrose, R. 2019a. Modeling Interwell Fracture Interference and Huff-N-Puff Pressure Containment in Eagle Ford Using EDFM. Paper SPE 195240, presented at SPE Oklahoma City Oil and Gas Symposium, Oklahoma City, Oklahoma, 9-10 April.

Fiallos Torres, M.X., Yu, W., Ganjdanesh, R., Kerr, E., Sepehrnoori, K., Miao, J., Ambrose, R. 2019b. Modeling Interwell Interference Due to Complex Fracture Hits in Eagle Ford Using EDFM. Paper IPTC 19468, presented at International Petroleum Technology Conference, Beijing, China, 26-28 March.

Fisher, M.K., Heinze, J.R., Harris, C.D., Davidson, B.M., Wright, C.A. and Dunn, K.P. 2004. Optimizing Horizontal Completion Techniques in the Barnett Shale Using Microseismic Fracture Mapping. Paper SPE 90051 presented at the 2004 SPE Annual Technical Conference and Exhibition, Houston, Texas, 26-29 September.

Gao G., Jiang H., Hagen P., Vink J., and Wells T., 2017. A Gauss-Newton Trust-Region Solver for Large-Scale History-Matching Problems. *SPE Journal* **22** (6): 1999-2011

Goodwin, N., 2015. Bridging the Gap Between Deterministic and Probabilistic Uncertainty Quantification Using Advanced Proxy Based Methods. Paper SPE 173301, presented at the SPE Reservoir Simulation Symposium, Houston, Texas, 23-25 February.

Guo, G., Evans, R. D., and Chang, M. M., 1994. Pressure-Transient Behavior for a Horizontal Well Intersecting Multiple Random Discrete Fractures. Paper SPE 28390, presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, 25-28 September.

Hagan, M.T., and Menhaj, M., 1994. Training feed-forward networks with the Marquardt algorithm. *IEEE Transactions on Neural Networks* **5** (6), 989-993.

Hastings, W., 1970. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometricka* **57** (1): 97-109.

Javadpour, F., Fisher, D., and Unsworth, M., 2007. Nanoscale Gas Flow in Shale Gas Sediments. *Journal of Canadian Petroleum Technology* **46** (10): 55-61.

Javadpour, F., 2009. Nanopores and Apparent Permeability of Gas Flow in Mudrocks (Shales and Siltstone). *Journal of Canadian Petroleum Technology* **48** (8): 16-21.

Luo, H., Mahiya, G., Pannett, S., and Benham, P. H., 2014. The Use of Rate-Transient-Analysis Modeling To Quantify Uncertainties in Commingled Tight Gas Production-Forecasting and Decline-Analysis Parameters in the Alberta Deep Basin. *SPE Reservoir Evaluation & Engineering* **17** (2): 209-219.

Maucec, M., Douma, S., Hohl, D., Lequijt, J., Jimenez, E. A., and Datta-Gupta, A., 2007. Streamline-Based History Matching and Uncertainty: Markov-chain Monte Carlo Study of an Offshore Turbidite Oil Field. Paper SPE 109943, presented at the SPE Annual Technical Conference and Exhibition, Anahelm, California, 11-14 November.

Maxwell, S.C., Kresse, O. and Rutledge, J. 2013. Modeling Microseismic Hydraulic Fracturing Deformation. Paper SPE 166312, presented at the SPE Annual Technical Conference and Exhibition, New Orleans, LA, 30 September – 2 October.

Mayerhofer, M. J., Lolon, E., Warpinski, N. R., Cipolla, C. L., Walser, D. W., and Rightmire, C. M. 2010. What Is Stimulated Reservoir Volume?. SPE Journal **25** (1): 89-98.

Metropolis, N., Rosenbluth, A.M., Rosenbluth, M.N., and Teller, A.H., 1953. Equation of State Calculations by Fast Computer Machines. *The Journal of Chemical Physics* **21** (6): 1087-1092.

Mirzabozorg A., Nghiem L., Chen Z., and Yang C., 2013. Differential Evolution for Assisted History Matching Process: SAGD Case Study. Paper SPE 165491, presented at the SPE Heavy Oil Conference – Calgary, Alberta, Canada, 11- 13 June.

Mohaghegh S. D., 2017. *Shale Analytics Data-Driven Analytics in Unconventional Resources*, Springer International Publishing.

Moinfar, A., Varavei, A., Sepehrnoori, K., and Johns, R. T., 2013. Development of a Coupled Dual Continuum and Discrete Fracture Model for the Simulation of Unconventional Reservoirs. Paper SPE 163647, presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, 18-20 February.

Moinfar, A., Varavei, A., Sepehrnoori, K., and Johns R.T., 2014. Development of an Efficient Embedded Discrete Fracture Model for 3D Compositional Reservoir Simulation in Fractured Reservoirs. *SPE Journal* **19** (2): 289-303.

Moinfar, A., Erdle, J. C., and Patel, K., 2016. Comparison of Numerical vs Analytical Models for EUR Calculation and Optimization in Unconventional Reservoirs. Paper SPE 180974, presented at the SPE Argentina Exploration and Production of Unconventional Resources Symposium, Buenos Aires, Argentina, 1-3 June.

Nejadi, S., Leung, J. Y. W., Trivedi, J. J., & Virues, C. J. J., 2015. Integrated Characterization of Hydraulically Fractured Shale Gas Reservoirs Production History Matching. Paper SPE 171664, presented at the SPE/CSUR Unconventional Resources Conference – Canada, Calgary, Alberta, Canada, 30 September–2 October.

Novlesky, A., Kumar, A., and Merkle, S., 2011. Shale Gas Modeling Workflow: From Microseismic to Simulation -- A Horn River Case Study. Paper SPE 148710, presented at the Canadian Unconventional Resources Conference, Calgary, Alberta, 15-17 November.

Oliver, D. and Chen, Y., 2011, Recent progress on reservoir history matching: A review, *Computational Geoscience* **15** (1), 185-221.

Peaceman, D. W., 1983. Interpretation of Well-Block Pressures in Numerical Reservoir Simulation With Nonsquare Grid Blocks and Anisotropic Permeability. *SPE Journal* **23** (03): 531-543.

Raghavan, R., Chen, C. C., and Agarwal, B., 1994. An Analysis of Horizontal Wells Intercepted by Multiple Fractures. Paper PETSOC-HWC-94-39, presented at the SPE/CIM/CANMET International Conference on Recent Advances in Horizontal Well Applications, March 20 - 23, Calgary, Canada

Rubin, B., 2010. Accurate Simulation of Non Darcy Flow in Stimulated Fractured Shale Reservoirs. Society of Petroleum Engineers. Paper SPE 132093, presented at the SPE Western Regional Meeting, Anaheim, California, 27-29 May.

Shabro, V., Torres-Verdin, C., and Javadpour, F., 2011. Numerical Simulation of Shale-Gas Production: From Pore-Scale Modeling of Slip-Flow, Knudsen Diffusion, and Langmuir Desorption to Reservoir Modeling of Compressible Fluid. Paper SPE 144355, presented at the North American Unconventional Gas Conference and Exhibition, The Woodlands, Texas, 14-16 June

Shabro, V., Torres-Verdin, C., and Sepehrnoori, K., 2012. Forecasting Gas Production in Organic Shale with the Combined Numerical Simulation of Gas Diffusion in

Kerogen, Langmuir Desorption from Kerogen Surfaces, and Advection in Nanopores. Paper SPE 159250, presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, 8-10 October.

Shaffner J., and Cheng. A, 2011. The Advantage of Incorporating Microseismic Data into Fracture Models. Paper CSUG/SPE 148780 presented at the Canadian Unconventional Resources Conference, Calgary, Alberta, Canada, 15-17 November.

Shahkarami, A., Mohaghegh S. D., and Hajizadeh, Y., 2015. Assisted History Matching Using Pattern Recognition Technology. Paper SPE 173405, presented at the SPE Digital Energy Conference and Exhibition, Woodlands, Texas, 3-5 March.

Shakiba, M., and Sepehrnoori, K., 2015. Using Embedded Discrete Fracture Model (EDFM) and Microseismic Monitoring Data to Characterize the Complex Hydraulic Fracture Networks. Paper SPE 175142, presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 28-30 September.

Sharma M.M., and Gadde P.B., 2004. Slick Water and Hybrid Fracs in the Bossier: Some Lessons Learnt. Paper SPE 89876, presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 26-29 September.

Shrivastava, K., Hwang, J., and Sharma., M., 2018. Formation of Complex Fracture Networks in the Wolfcamp Shale: Calibrating Model Predictions with Core Measurements from the Hydraulic Fracturing Test Site. Paper SPE 191630, presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas, 24-26 September.

Slotte, P. A., and Smorgrav, E., 2008. Response Surface Methodology Approach for History Matching and Uncertainty Assessment of Reservoir Simulation Models. Society of Petroleum Engineers. Paper SPE 113390, presented at the Europec/EAGE Conference and Exhibition, Rome, Italy, 9-12 June.

Stalgorova, K., and Mattar, L., 2013. Analytical Model for Unconventional Multifractured Composite Systems. *SPE Reservoir Evaluation & Engineering* **16** (3): 246-256

Stat-Ease Inc. 2017. Design Expert Software (DX11). Minneapolis, Minnesota: Stat-Ease, Inc.

Taleghani, A. D., and Olson, J. E., 2013. How Natural Fractures Could Affect Hydraulic-Fracture Geometry. *SPE Journal* **19** (1): 161-171.

Tripoppoom, S., Yu, W., Huang, H., Sepehrnoori, K., Song, W., and Dachanuwattana, S., 2019. A Practical and Efficient Iterative History Matching Workflow for Shale Gas Well Coupling Multiple Objective Functions, Multiple Proxy-based MCMC and EDFM. *Journal of Petroleum Science and Engineering* **176**: 594–611.

Vazquez, O., Young, C., Demyanov, V., Arnold, D., Fisher, A., MacMillan, A., and Christie, M., 2015. Produced-Water-Chemistry History Matching in the Janice Field. *SPE Reservoir Evaluation & Engineering* **18** (4): 564-576

Wantawin, M., Yu, W., and Sepehrnoori, K., 2017a. An Iterative Response-Surface Methodology by Use of High-Degree-Polynomial Proxy Models for Integrated History Matching and Probabilistic Forecasting Applied to Shale-Gas Reservoirs. *SPE Journal* **22** (6): 2012-2031.

Wantawin, M., Yu, W., and Sepehrnoori, K., 2017b. An Iterative Work Flow for History Matching by Use of Design of Experiment, Response-Surface Methodology, and Markov Chain Monte Carlo Algorithm Applied to Tight Oil Reservoirs. *SPE Reservoir Evaluation & Engineering* **20** (3): 613-626.

Warpinski, N.R., Mayerhofer, M.J., Vincent, M.C., Cipolla, C.L., and Lolon, E.P. 2008. Stimulating Unconventional Reservoirs: Maximizing Network Growth while Optimizing Fracture Conductivity. Paper SPE 114173 presented at the 2008 SPE Unconventional Reservoirs Conference, Keystone, Colorado, 10-12 February.

Wu, K., and Olson, J.E., 2015. Simultaneous Multifracture Treatments: Fully Coupled Fluid Flow and Fracture Mechanics for Horizontal Wells. *SPE Journal* **20** (2): 337-346.

Wu, K., and Olson, J.E., 2016. Numerical Investigation of Complex Fracture Networks in Naturally Fractured Reservoirs. *SPE Production & Operations* **31** (4): 300-309.

Xie, J., Yang, C., Gupta, N., King, M. J., and Datta-Gupta, A., 2015. Integration of Shale-Gas-Production Data and Microseismic for Fracture and Reservoir Properties With the Fast Marching Method. *SPE Journal* **20** (2): 347-359.

Xu W., Calvez J., and Thiercelin M, 2009. Characterization of Hydraulically-Induced Fracture Network Using Treatment and Microseismic Data in a Tight-Gas Formation: A Geomechanical Approach. Paper SPE 125237 presented at the SPE Tight Gas Completions Conference, San Antoni, Texas, 15-17 June.

Xu, Y., Cavalcante Filho, J.S.A., Yu, W., and Sepehrnoori, K., 2017a. Discrete-Fracture Modeling of Complex Hydraulic-Fracture Geometries in Reservoir Simulators. *SPE Reservoir Evaluation & Engineering* **20** (2): 403-422.

Xu, Y., Yu, W., and Sepehrnoori, K., 2017b. Modeling Dynamic Behaviors of Complex Fractures in Conventional Reservoir Simulators. Paper URTeC 2670513, presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Austin, Texas, 24-26 July.

Xu, Y., Yu, W., Li, N., Lolon, E., and Sepehrnoori, K., 2018. Modeling Well Performance in Piceance Basin Niobrara Formation using Embedded Discrete Fracture Model. Paper URTeC 2901327, presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, Texas, 23-25 July.

Yang, C., Nghiem, L., Erdle, J., Moinfar, A., Fedutenko, E., Li, H., Mirzabozorg, A., and Card, C., 2015. An Efficient and Practical Workflow for Probabilistic Forecasting of Brown Fields Constrained by Historical Data. Paper SPE 175122, presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 28-30 September.

Yeh, T., Uvieghara, T., Jennings, J. W., Chen, C., Alpak F. O., and Tendo, F., 2016. A Practical Workflow for Probabilistic History Matching and Forecast Uncertainty Quantification: Application to a Deepwater West Africa Reservoir. Paper SPE 181639, presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, 25-28 September.

Yin, J., Xie, J., Datta-Gupta, A., and Hill, A. D., 2011. Improved Characterization and Performance Assessment of Shale Gas Wells by Integrating Stimulated Reservoir Volume and Production Data. Paper SPE 148969, presented at the SPE Eastern Regional Meeting, Columbus, Ohio, 17-19 August.

Yu, W., and Sepehrnoori, K., 2014. Simulation of Gas Desorption and Geomechanics Effects for Unconventional Gas Reservoirs. *Fuel* **116**: 455-464.

Yu, W., Sepehrnoori, K., and Patzek, T.W., 2016. Modeling Gas Adsorption in Marcellus Shale with Langmuir and BET Isotherms. *SPE Journal* **21** (2): 589-600.

Yu, W., and Sepehrnoori, K. 2018. *Shale Gas and Tight Oil Reservoir Simulation*, 1st Ed.; Publisher: Elsevier, Cambridge, USA. ISBN: 978-0-12-813868-7.

Yu, W., Xu, Y., Weijermars, R., Wu, K., and Sepehrnoori, K. 2018a. A Numerical Model for Simulating Pressure Response of Well Interference and Well Performance in Tight Oil Reservoirs with Complex-Fracture Geometries Using the Fast Embedded-Discrete-Fracture-Model Method. *SPE Reservoir Evaluation & Engineering* **2** (21): 489-502.

Yu, W., Xu, Y., Liu, M., Wu, K., and Sepehrnoori, K., 2018b. Simulation of Shale Gas Transport and Production with Complex Fractures using Embedded Discrete Fracture Model. *AIChE Journal* **64** (6): 2251-2264.

Yu, W., Tripoppoom, S., Sepehrnoori, K., and Miao, J., 2018c. An Automatic History-Matching Workflow for Unconventional Reservoirs Coupling MCMC and Non-Intrusive EDFM Methods. Paper SPE-191473-MS, presented at the SPE Annual Technical Conference and Exhibition, 24-26 September, Dallas, Texas.

Yu, W., Wu, K., Liu, M., Sepehrnoori, K., and Miao, J. 2018d. Production Forecasting for Shale Gas Reservoirs with Nanopores and Complex Fracture Geometries Using An Innovative Non-Intrusive EDFM Method. Paper SPE 191666, presented at SPE Annual Technical Conference and Exhibition, Dallas, Texas, 24-26 September.

Yu, W., Hu X., Liu, M. and Wang, W., 2019a. Investigation of the Effect of Natural Fractures on Multiple Shale-Gas Well Performance Using Non-Intrusive EDFM Technology. *Energies* **12** (5): 932.

Yu, W., Zhang, Y., Varavei, A., Sepehrnoori, K., Zhang, T., Wu, K., and Miao, J. 2019b. Compositional Simulation of CO2 Huff-n-Puff in Eagle Ford Tight Oil Reservoirs with CO2 Molecular Diffusion, Nanopore Confinement and Complex Natural Fractures. *SPE Reservoir Evaluation & Engineering*, in preprint.

Zhang, Z., and Fassihi, M. R., 2013. Uncertainty Analysis and Assisted History Matching Workflow in Shale Oil Reservoirs. Paper URTeC 1581398, presented at the Unconventional Resources Technology Conference, Denver, Colorado, 12-14 August.